```
REMARK      Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            On Aug 19, 2003 this sequence version replaced gi:15530217.
COMMENT     Contact: MGC help desk
            Email: cgapbb-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 14 Row: m Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 34222363.
FEATURES             Location/Qualifiers
     source          1..2096
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:11168 IMAGE:3842949"
                     /tissue_type="Placenta; choriocarcinoma"
                     /clone_lib="NIH_MGC_21"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1..2096
                     /gene="BAIAP2L1"
                     /note="synonym: IRTKS"
                     /db_xref="GeneID:55971"
     CDS             217..1752
                     /gene="BAIAP2L1"
                     /codon_start=1
                     /product="BAI1-associated protein 2-like 1"
                     /protein_id="AAH13888.1"
                     /db_xref="GI:15530218"
                     /db_xref="GeneID:55971"
                     /translation="MSRGPEEVNRLRTESTYRNVMEQFNPGLRNLINLGNYEKAVNAM
                     ILAGKAYIDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNESLDENFKKFHKEIIH
                     ELEKKIELDVKVNNATLKRYQTEHKNKLESLEKSQAELKIRKKSQGSRNALKYEHKE
                     IEYVETVTRSQSBIQKFIADGCKEALLEEKRRFCPLVDKHGFANHIHYHLQSAELL
                     NSKLPRWQTCVDAIKVPEKIMNMIEEIKTPASTPVSGTPQASPMIERSNVVRKDYDT
                     LSKCSPKMPPASPGRAYTSPLIDMFNANPATAAPNSQRVNNSTSTSEDSLQRSVSVAT
                     GLNMMKQQKVKIFPHTAGSNKTLLSFAQGDVITLLIPEEKDGWLYGEHDVSKARGWF
                     PSSYTKLLEENETEAVTVPTPSPTPVRSISTVNLSENSSVVIPPDYLECLSMGAAAD
                     RRADSARTTSTFKAPASKPETAAPNDANGTAKPPFLSGENPFATVKLRPTVTNDRSAP
                     IIR"
ORIGIN
```

```
        Query Match           100.0%;    Score 2096;    DB 8;    Length 2096;
        Best Local Similarity 100.0%;    Pred. No. 0;
        Matches 2096;  Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy        1 CTCCCTGCTCTCCTCGAAGAAGGCCAGGGCGGGGGGCTCGCCGCAAGTGTTTGACATTTCGCAG   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 CTCCCTGCTCTCCTCGAAGAAGGCCAGGGCGGGGGGCTCGCCGCAAGTGTTTGACATTTCGCAG   60

Qy       61 CGGAGACGCGCGCGCGGGCCCACTCTCGGCGGGCCGGGGGGGGGGGGGGGGGAGGGGGAGAGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CGGAGACGCGCGCGCGGGCCCACTCTCGGCGGGCCGGGGGGGGGGGGGGGGGAGGGGGAGAGC  120

Qy      121 CCCTTAGTGCGGCGCCCCCCGCGCCCTCCCGCGTCGGCGCCGCCAGAGGGCCCAGCCG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CCCTTAGTGCGGCGCCCCCCGCGCCCTCCCGCGTCGGCGCCGCCAGAGGGCCCAGCCG  180

Qy      181 CGTCCTCCCCGCGCTTGCTGCGCCCCGGCCCATGTCCCGGGGGCCGCAGCCATGTCCCGGAGGAGGTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CGTCCTCCCCGCGCTTGCTGCGCCCCGGCCCATGTCCCGGGGGCCGCAGCCATGTCCCGGAGGAGGTG  240

Qy      241 AACCGGCTCACGGAGGAGCACCTACCGGAGAGCACCTACCGGAGAATGTTATGTGAACAGTTCAATCCTGGGGCTGGA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AACCGGCTCACGGAGGAGCACCTACCGGAGAGCACCTACCGGAGAATGTTATGTGAACAGTTCAATCCTGGGGCTGGA  300

Qy      301 AATTTAATAAACCTGGGGGAAAAATTATGAGAAAGCTGTAAACCTATGATCCTGGCAGGA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AATTTAATAAACCTGGGGGAAAAATTATGAGAAAGCTGTAAACCTATGATCCTGGCAGGA  360

Qy      361 AAAGCCTACTACGGATGGAGTGGGCCAAGATCGGTGAGATTGGTGAGATTGGTGAGATTGTTGGGTTGTCA  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAAGCCTACTACGGATGGAGTGGGCCAAGATCGGTGAGATTGGTGAGATTGTTGGGTTGTCA  420

Qy      421 ACTGAACTGGGACATGTGCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 ACTGAACTGGGACATGTGCCTCATAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT  480

Qy      481 CTTGATGAAAATTTAAAAAATTCCACAAAGAGAGATTATCCATGAGCTGGAGAAGAAGATA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CTTGATGAAAATTTAAAAAAATTCCACAAAGAGAGATTATCCATGAGCTGGAGAAGAAGATA  540

Qy      541 GAACTTGACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCCAAAACAGAACAAGAAT  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GAACTTGACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCCAAAACAGAACAAGAAT  600

Qy      601 AAATTAGAGTCTTTGGAGAAGATCAGAGTTGAAGAAGATCAGAGGAAAAGCCAA  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AAATTAGAGTCTTTGGAGAAGATCAGAGTTGAAGAAGATCAGAGGAAAAGCCAA  660

Qy      661 GGAAGCCGAAACGCACTCAAATAATGAACACAAAAGAAATTGAGTAGTGTCGGAGACCGTTACT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GGAAGCCGAAACGCACTCAAATAATGAACACAAAAGAAATTGAGTAGTGTCGGAGACCGTTACT  720

Qy      721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGAAGAGGCTCTGCTTGAA  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGAAGAGGCTCTGCTTGAA  780

Qy      781 GAGAAGAGGCGCTTCTCGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GAGAAGAGGCGCTTCTCGCTTTCTGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAT  840

Qy      841 TATTATCACTTACAGTCTGCGAGAACTACTGAATTCCAAGCTCGCCTCGGTGGCAGGAGACC  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TATTATCACTTACAGTCTGCGAGAACTACTGAATTCCAAGCTCGCCTCGGTGGCAGGAGACC  900

Qy      901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACC  960

Qy      961 CCAGCCTCTACCCCGTGTCTGGAACTCCCTCAGGCCTTCACCCATGATCGAGAGGAAGCAAT  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CCAGCCTCTACCCCGTGTCTGGAACTCCCTCAGGCCTTCACCCATGATCGAGAGGAAGCAAT  1020

Qy     1021 GTGGTTAGGGAAAGATTACGGACACCCTTTCTAAAATGCTTCACCAAAGATGCCCCCGCTTCCT  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GTGGTTAGGGAAAGATTACGGACACCCTTTCTAAAATGCTTCACCAAAGATGCCCCCGCTTCCT  1080

Qy     1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGAATATGTTAATTAACCCAGTCACGGCTGCC  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TCAGGCAGAGCATATACCAGTCCCTTGATCGAATATGTTAATTAACCCAGTCACGGCTGCC  1140

Qy     1141 CCGGAATTCACAAAGGGTAAATATTCAACAGTGACTACTTCCGAAGATCCCAGTTTACAGCGA  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CCGGAATTCACAAAGGGTAAATATTCAACAGTGACTACTTCCGAAGATCCCAGTTTACAGCGA  1200

Qy     1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAGACAGAAAAGTGAAGGCCATCTTC  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAGACAGAAAAGTGAAGGCCATCTTC  1260

Qy     1261 CCGCACACTGCGGGGGCTCCACACAAGACCTTACTCAGCTTTGCACAGGGAGAGTCATCACG  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CCGCACACTGCGGGGGCTCCACACAAGACCTTACTCAGCTTTGCACAGGGAGAGTCATCACG  1320

Qy     1321 CTGCTCCATCCCGAGGAGAGGATGGCGTGGCTCTATGGAGAACACGACGTGTCCAAGGCG  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CTGCTCCATCCCGAGGAGAGGATGGCGTGGCTCTATGGAGAACACGACGTGTCCAAGGCG  1380
```

```
OM nucleic - nucleic search, using sw model

                    Copyright (c) 1993 - 2006  Bioacceleration Ltd.
                    GenCore version 5.1.7

Run on:             March 20, 2006, 13:39:55 ; Search time:10285 seconds
                                                (without alignments)
                                                11584.237 Million cell updates/sec

Title:              US-10-801-292-1
Perfect score:      2096
Sequence:           1 ctcctctgctcctcgaagaa........aaaaaaaaaaaaaaaaaaaa 2096
Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0
Searched:           5883141 seqs, 2842172553 residues
Total number of hits satisfying chosen parameters:            11766282
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing:    Minimum Match 0%
                    Maximum Match 10%
                    Listing first 45 summaries

Database :          GenEmbl:*
                    1: gb_ba:*
                    2: gb_in:*
                    3: gb_env:*
                    4: gb_om:*
                    5: gb_ov:*
                    6: gb_pat:*
                    7: gb_pat:*
                    8: gb_pr:*
                    9: gb_ro:*
                    10: gb_sts:*
                    11: gb_sy:*
                    12: gb_un:*
                    13: gb_vi:*
                    14: gb_htg:*
                    15: gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	2096	100.0	2096	8	BC013888	Homo sapi
2	2025.2	96.6	2527	8	AP119666	Homo sapi
3	1821.4	86.9	2317	6	AX058219	Sequence
4	1560	74.4	1580	8	AK027109	Homo sapi
5	1235.4	58.9	1491	6	CQ730103	Sequence
6	1163	55.5	3040	9	BC015459	Mus muscu
7	1042.8	49.8	2141	9	BC089216	Rattus no
8	859	41.0	2521	9	AK027142	Homo sapi
9	745.2	35.6	2064	5	BX933564	Gallus ga
10	671	32.0	2296	8	AK074419	Homo sapi
11	594.4	28.4	3616	6	BC061676	Xenopus l
12	556	26.5	559	6	AX379210	Sequence
13	397	18.9	685	5	BX931244	Gallus ga
14	382.6	18.3	2563	5	BC068330	Danio rer
15	381	18.2	2566	5	AY398378	Danio rer
16	377.8	18.0	2570	5	BC050238	Danio rer
17	365.8	17.5	2603	5	BC092786	Danio rer
c 18	357	17.0	11840	8	AC093169	Homo sapi
19	355	16.9	3128	6	CQ843301	Sequence
20	355	16.9	3128	8	AK124269	Homo sapi
21	307.2	14.7	1678	9	AF390179	Mus muscu
22	285.2	13.6	490	5	BX930239	Gallus ga
23	266.6	12.7	2078	9	CGU41899	Cricetinae
c 24	265.4	12.7	132072	8	AC004841	Homo sapi
c 25	265.6	12.7	156195	8	AC093799	Homo sapi
25	265.6	12.6	2111	9	AY037934	Rattus no
26	265	12.6	1798	9	BC074009	Sequence
26	265.6	12.5	2548	9	AF390178	Mus muscu
27	261.6	12.5	2545	9	BC006620	Mus muscu
28	261.6	12.5	1798	9	BC016411	Mus muscu
29	261.6	12.4	2080	9	AR069610	Sequence
30	261.6	12.4	2080	6	AR352563	Sequence
31	259.6	12.4	2080	6	AR352563	Sequence
32	259.6	12.4	1078	6	U70669	Human Fas-1
33	259.6	12.3	2033	6	AB017120	Homo sapi
34	258	12.3	2072	8	BC014020	Homo sapi
35	258	12.3	2077	8	AB169737	Macaca fa
36	258	12.3	2129	8	AB169737	Macaca fa
37	258	12.3	2129	6	CQ896663	Sequence
38	258	12.3	2137	6	BD226257	Pancreati
39	258	12.3	2137	8	AB015020	Homo sapi
40	258	12.3	2519	6	AX011646	Sequence
41	258	12.3	2519	6	BC032559	Homo sapi
42	258	12.3	2877	8	BC032559	Homo sapi
43	258	12.3	3168	8	AB017119	Homo sapi
44	258	12.3	3168	8	AB015019	Homo sapi
45	256.4	12.2	1910	8	AK222670	Homo sapi

ALIGNMENTS

```
RESULT 1
BC013888
LOCUS       BC013888               2096 bp    mRNA    linear   PRI 15-FEB-2005
DEFINITION  Homo sapiens BAI1-associated protein 2-like 1, mRNA (cDNA clone
            MGC:11168 IMAGE:3842949), complete cds.
ACCESSION   BC013888
VERSION     BC013888.2  GI:33871229
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2096)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S, Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M, Gay,L.J Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2096)
  AUTHORS   .
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
```

LOCUS   AF119666   2527 bp   mRNA   linear   PRI 26-AUG-2003
DEFINITION   Homo sapiens insulin receptor tyrosine kinase substrate mRNA,
             complete cds.
ACCESSION   AF119666
VERSION   AF119666.2 GI:34223696
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE   1 (bases 1 to 2527)
  AUTHORS   Peng,Y., Huang,C., Wu,T., Huang,Q., Gu,Y., Zhang,L., Li,Y., Xu,S.,
             Fu,G., Wang,Y., Chen,Z. and Han,Z.
  TITLE   A novel gene expressed in the human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2527)
  AUTHORS   Li,Y., Han,Z., Zhang,C., Wang,Y., Chen,Z. and Fu,G.,
  TITLE   Direct Submission
  JOURNAL   Submitted (26-AUG-2003) Chinese National Human Genome Center at
             Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong,
             Shanghai 201203, P. R. China
REFERENCE   3 (bases 1 to 2527)
  AUTHORS   Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
             Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-JAN-1999) Chinese National Human Genome Center at
             Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong,
             Shanghai 201203, P. R. China
  REMARK   Sequence update by submitter
COMMENT   On Aug 26, 2003 this sequence version replaced gi:6563257.
FEATURES   Location/Qualifiers
     source   1..2527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     CDS   217..1752
                     /note="IRTKS"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="insulin receptor tyrosine kinase substrate"
                     /protein_id="AAR17223.2"
                     /db_xref="GI:34223697"
                     /translation="MSRGPEEVNRLTESTYRNVMEQNPGLRNLINLGKNYEKAVNAM
                     ILAGKAYYDGVAKIGETATGSPVSTELGHVLIEISSTHKKLNESLLDENFKKFHKEIIH
                     ELEKKIELDVKYMNATLKRYQTEHRKNKLESLEKSQAELKIRRKSQGSRNALKYEHKE
                     IEYVETVTSRQSEIQKPIADGCKEALLEEKRRFCFLVDKHGCFPANHIHYYHLQSABLL
                     NSKLPRWQETCVDAIKVPEKIMNMIEEIKTPASTPVSGTPQASPMIERSNVRKDYDT
                     LSKCSFKMPPAPSGRAYTSPLIDMFENNPATAAPNSQRVNNSTGTSEDPSLQRSVSVAT
                     GLNMMKKQKVKIIFPHFAGSNKTLLSFAQGDVITLLIPEEKDGWLYGEHDVSKARGWF
                     PSSYTKLLEENETEAVTVPTPSSPTPVRSISTVNLSENSSVVTPPPDYLECLSMGAAD
                     RRADSARTISTFKAPASKPETAAPNDANGTAKPFFLSGENPFATVKLRPTVTNDRSAP
                     IIR"
ORIGIN

RESULT 2

AF119666

Query Match   96.6%; Score 2025.2; DB 8; Length 2527;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 2027; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGGCGGGGGGGCTGCCCGGCCAAGTTTTGACATTTTCGCAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CTCCTCTGCTCCTCGAAGAAGGCCAGGGGCGGGGGGGCTGCCCGGCCAAGTTTTGACATTTTCGCAG 60

Qy   61 CGGAGGACGGCGGCGGCGGCGGCACTCTCGGGGACGGGCTGCGCCGGCCAAGCCCTTCCAGAGC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGGAGGACGGCGGCGGCGGCGGCACTCTCGGGGACGGGCTGCGCCGGCCAAGCCCTTCCAGAGC 120

Qy   121 CCCTTAGTCGCGGCGCGGCGCTTGCTCGCGCCGCGGCCGCCCGGCGGGGGCCCGAGGAGGTG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCCTTAGTCGCGGCGCGGCGCTTGCTCGCGCCGCGGCCGCCCGGCGGGGGCCCGAGGAGGTG 180

Qy   181 CGTCCTCCCGGCGGCGCTTGCTCGCGCCGCGGCCGCCCGGCGGGGGCCGCGAGGAGGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CGTCCTCCCGGCGGCGCTTGCTCGCGCCGCGGCCGCCCGGCGGGGGCCGCGAGGAGGTG 240

Qy   241 AACCGGGCTCACCGAGAGACACTTACCCGGAATGTTATGGGAACAGTTCAATCCTGGGCTGCGA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AACCGGGCTCACCGAGAGACACTTACCCGGAATGTTATGGGAACAGTTCAATCCTGGGCTGCGA 300

Qy   301 AATTTAATAAACCTGGGGAAAAAATTATGAGAAAAGCTGTAAACGCTATGGATCCTGGCCAGGA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AATTTAATAAACCTGGGGAAAAAATTATGAGAAAAGCTGTAAACGCTATGGATCCTGGCCAGGA 360

Qy   361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCA 420

Qy   421 ACTGAACTGGGACACATGTCCTCATTAGAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ACTGAACTGGGACACATGTCCTCATTAGAGAGATTTCAAGTACCCACAAGAAACTCAACGAGAGT 480

Db   1321 CTGCTCATCCCGAGGAGGAAGGAGATGGCTGGCTTATGGAGGAACACGACGACGTGTCCAAGGCG 1380
Qy   1381 AGGGGTTGGTTCCCGTCGTGTGGTCGTACACGAGAGTTGCTGGAAGAAAAATGAGACAGAGCAGTG 1440
Db   1381 AGGGGTTGGTTCCCGTCGTGTGGTCGTACACGAGAGTTGCTGGAAGAAAAATGAGACAGAGCAGTG 1440
Qy   1441 ACCGTGCCCACGCCACCAGCCCCACACCAGTGAGAGCATCAGCACGTGAAACTTGTCTGAG 1500
Db   1441 ACCGTGCCCACGCCAAGCCCCACACCAGTGAGAGCATCAGCACGTGAAACTTGTCTGAG 1500
Qy   1501 AATAGACAGTGTTGTTCATCCCCCCGACTACTTGGAATGCTTGTCCATGGGGGGCAGCT 1560
Db   1501 AATAGACAGTGTTGTTCATCCCCCCGACTACTTGGAATGCTTGTCCATGGGGGGCAGCT 1560
Qy   1561 GCCGACAGGAGGAGCAGGATTCGGCGCCAGGACGACATCCACCTTTAAGGCCCCAGGGGTTCCAAG 1620
Db   1561 GCCGACAGGAGGAGCAGGATTCGGCGCCAGGACGACATCCACCTTTAAGGCCCCAGGGGTTCCAAG 1620
Qy   1621 CCCGAGACCGCGGCGCGCGGCTCCTAACGATGCCAAGCGGCCCTTTTCTCAGCGGA 1680
Db   1621 CCCGAGACCGCGGCGCGCGGCTCCTAACGATGCCAAGCGGCCCTTTTCTCAGCGGA 1680
Qy   1681 GAAAAACCCTTTGCCACTGTGAAACTCCGGCCCGGCACTGTGACGAATGATCGGCCCACCC 1740
Db   1681 GAAAAACCCTTTGCCACTGTGAAACTCCGGCCCGGCACTGTGACGAATGATCGGCCCACCC 1740
Qy   1741 ATCATTCGATGAGGAGGACAGCAGCCAAGGACTCTCCCGGGGCCCTTCTCCCGGTTCTCCCCTTGCGGA 1800
Db   1741 ATCATTCGATGAGGAGGACAGCAGCCAAGGACTCTCCCGGGGCCCTTCTCCCGGTTCTCCCCTTGCGGA 1800
Qy   1801 ATGATCGGGCGGCCATCCTCTCTGCCACGTGCTGCGACGGTCGGGGAAGCTTCAGTGAGAGGCCT 1860
Db   1801 ATGATCGGGCGGCCATCCTCTCTGCCACGTGCTGCGACGGTCGGGGAAGCTTCAGTGAGAGGCCT 1860
Qy   1861 AACTCTAAATGTCGCCTGCTTAAGCAAAATCATCGCTTCTCTGTTCGTTCGTGTGGGTTGAC 1920
Db   1861 AACTCTAAATGTCGCCTGCTTAAGCAAAATCATCGCTTCTCTGTTCGTTCGTGTGGGTTGAC 1920
Qy   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTTCTAATGACCAGCCATTTAAAAT 1980
Db   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTTCTAATGACCAGCCATTTAAAAT 1980
Qy   1981 ATTTTCTTCCTATTCTCTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAG 2040
Db   1981 ATTTTCTTCCTATTCTCTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAG 2040
Qy   2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
Db   2041 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096

```
Qy    481 CTTGATGAAAATTTTAAAAAATTCCACAAGAGAGATTATCCATGAGCTTGGAGGAAGAGATA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGATGAAAATTTTAAAAAATTCCACAAGAGAGATTATCCATGAGCTGGAGGAAGAGATA 540

Qy    541 GAAACTTGACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCACAGACAACAAGAAT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAAACTTGACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCACAGACAACAAGAAT 600

Qy    601 AAAATTAGAGTCTTTGGAGAAAATCCAAGCTGAGTTGAAGAGAAAAGCCAA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAAATTAGAGTCTTTGGAGAAAATCCAAGCTGAGTTGAAGAGAAAAGCCAA 660

Qy    661 GGAAGCCGAAAACGCACTCAAATATGAACACAAAGAAATTGAGTAATGCACCGGTTACT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGAAGCCGAAAACGCACTCAAATATGAACACAAAGAAATTGAGTAATGCACCGGTTACT 720

Qy    721 TCTCCGTCAGAGTGAAAATCCAGAAATTCAGAAAATTCATTGCGAGATGGTTGCTTGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCTCCGTCAGAGTGAAAATCCAGAAATTCAGAAAATTCATTGCGAGATGGTTGCTTGAA 780

Qy    781 GAGAAGAGGGCGCCTTCTGCTTTCTTGGTTGTTGATAAGCACTGTGGCTTTGCAAACCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAAGAGGGCGCCTTCTGCTTTCTTGGTTGTTGATAAGCACTGTGGCTTTGCAAACCA 840

Qy    841 TATTATCACTTACAGTCTGCAGAACTACTGGTTGAATCCAAGCTGCCTCGGGTGGCAGACC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TATTATCACTTACAGTCTGCAGAACTACTGGTTGAATCCAAGCTGCCTCGGGTGGCAGACC 900

Qy    901 TGTGTTGATGCCATCAAAAGTGCCAGAGAAATTCATGAATATGATGGAGAAAATAAAGACC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGTGTTGATGCCATCAAAAGTGCCAGAGAAATTCATGAATATGATGGAGAAAATAAAGACC 960

Qy    961 CCAGCCTCTACCCCGTGTCTGGAACTCCCAGGCTTCACCATGATCGAGAGAAGCAAT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCAGCCTCTACCCCGTGTCTGGAACTCCCAGGCTTCACCATGATCGAGAGAAGCAAT 1020

Qy   1021 GTGGTTAGGAAAGATTACGACACCCCTTTCTAAAATGCTCACCAAAGATGCCCCCGCTCCT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTGGTTAGGAAAGATTACGACACCCCTTTCTAAAATGCTCACCAAAGATGCCCCCGCTCCT 1080

Qy   1081 TCAGGCAGAGACATATACCAGTCCCTTGATGATGTTAATAACCCAGCCAGGCCTGCC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCAGGCAGAGCATATACCAGTCCCTTGATGATGTTAATAACCCAGCCAGGCCTGCC 1140

Qy   1141 CCGAATTCACAAAAGGGGTAAATAATTCAACAGGTAATAATTCAACAGTTTTACAGGGA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCGAATTCACAAAAGGGGTAAATAATTCAACAGGTAATAATTCAACAGTTTTACAGGGA 1200

Qy   1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGCAGAAGAGACAGAGAACCAATCTTC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGCAGAAGAGACAGAGAACCAATCTTC 1260

Qy   1261 CCGCACACTGGGGCTCCAACAAGACCTTACTTCACTTCAGCTTTGCACAGGGAGAGTCATCACG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCGCACACTGCGGGCTCCAACAAGACCTTACTCACTTTGCACAGGGAGAGTCATCACG 1320

Qy   1321 CTGCTCATCCCCGAGGAGAAGATGGCTGGGCTCTCATGAGAGAACACGAGCGG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGCTCATCCCCGAGGAGAAGATGGCTGGGCTCTCATGAGAGAACACGAGCGG 1380

Qy   1381 AGGGGTTGGGTTCCCGTGTCGTGCTACACGAGAAAATGAGACAGAAGCACTG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AGGGGTTGGGTTCCCGTGTCGTGCTACACGAGAAAATGAGACAGAAGCACTG 1440

Qy   1441 ACCGTGCCCCAGCCCACGCCCAGTGAGAAGCATCAGCAGCCGTGAAACTTGTCTGAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ACCGTGCCCCAGCCCACGCCCAGTGAGAAGCATCAGCAGCCGTGAAACTTGTCTGAG 1500

Qy   1501 AATAGCAGTGTTGTTGTTCATCCCCCACCCGACTACTTGGAATGGCTTATCCATGGGGGGGCAGCT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AATAGCAGTGTTGTTGTTCATCCCCCACCCGACTACTTGGAATGGCTTATCCATGGGGGGGCAGCT 1560

Qy   1561 GCCGCAGGAGGAGGCAGAGATTCGGCGGCAGGAGCGGCGACGACATCCACACCTTTTAAGGCCCCAGGGTCCAAG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGACAGGAGGAGCAGAGATTCGGCGGCAGGAGCGGCGACGACATCCACACCTTTTAAGGCCCCAGGGTCCAAG 1620

Qy   1621 CCCGAGACCGGCGCCTCCTTAACGATTGCCAAGCAGGAGCGGCAAACGGGACTGCAAAGCGCCCTTTTTCTCAGCCGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCCGAGACCGGCGCCTCCTTAACGATTGCCAAGCAGGAGCGGCAAACGGGACTGCAAAGCGCCCTTTTTCTCAGCCGGA 1680

Qy   1681 GAAAAACCCCTTTGCCACTGTGAAAACTCCGGCGACCTGTGACGAATGATGCTCGGCCACC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAAAAACCCCTTTGCCACTGTGAAAACTCCGGCGACCTGTGACGAATGATGCTCGGCCACC 1740

Qy   1741 ATCATTCGATGAGAGACAGCGCAAAGGACTCTCGCCCCCCCGGACACTGTGACGAATGATGCTCGGCCACC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATCATTCGATGAGAGACAGCGCAAAGGACTCTCGCCCCCCCGGACACTGTGACGAATGATGCTCGGCCACC 1800

Qy   1801 ATGATGGGGCGCCGCCATCCTGTCTCGCCGACGGTCGGGAAGGCTTCAGTGGGAGAGGGCCT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATGATGGGGCGCCATCCTGTCTCGCCGACGGTCGGGAAGGCTTCAGTGGGAGAGGGCCT 1860

Qy   1861 AACTCTAAGTGTGGCCTGCTTAAGCAAAATCATGCTTTCTTGTTTTCACGTAGTTGGGTTGAC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AACTCTAAATGTCGGCCTGCTTAAGCAAAATCATGCTTTCTTGTTTTCACGTAGTTGGGTTGAC 1920

Qy   1921 AAGTTTCTCGCCCTTTAAGATAAAATGAGTAATAGTCTTAATGACCAGCTCAGCCATTTAAAAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGTTTCTGCCCTTTAAGATAAAATGAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAAT 1980

Qy   1981 ATTTTTCTTCCTATTTCTGTTCAAGAAAACAGTAAACTTGGTTTCAATCTTTA 2030
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATTTTTCTTCCTATTTCTGTTCAAGAAAACAGTAAACTTGGTTTCAATCTTTA 2030
```

RESULT 3
AX058219

```
LOCUS       AX058219                2317 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 89 from Patent WO0077040.
ACCESSION   AX058219
VERSION     AX058219.1  GI:12310720
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Hillman,J.L., Lal,P., Bandman,O., Baughn,M.R.,
            Azimzai,Y., Yang,J., Reddy,R. and Lu,D.A.
  TITLE     Human intracellular signaling molecules
  JOURNAL   Patent: WO 0077040-A 89 21-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..2317
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match    86.9%;  Score 1821.4;  DB 6;  Length 2317;
Best Local Similarity   99.9%;
Matches 1822;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Qy    208 GCCGCAGGCCATGTCCGGCGGGGGCCCGGAGGAGGTGGAACCGGCGCTCACGGGAGCCACCTACCGG 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     30 GCCGCAGCCATGTCCGGCGGGGCCCGGAGGAGGTGGAACCGGCGCTCACGGGAGCCACCTACCGG 89

Qy    268 AATGTTATGGAACAGTTCAATCCTGGGGCTTGCGAAATTTAATAAACCTGGGGAAAATTTAT 327
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     90 AAGTTATGGAACAGTTCAATCCTGCGGCTTGCGAAATTTAATAAACCTGGGGAAAAATTAT 149

Qy    328 GAGAAAAGCTGTAAACGCTATGATGATCCTGGCGAGGAAAAAGCCTACTACGATGGAGTGGCCAAG 387
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    150 GAGAAAGTCGTAAACGCTATGATGATCCTGGCGAGGAAAAAGCCTACTACGATGGAGTGGCCAAG 209
```

```
Qy    388 ATCGGTGAGATTGCCACTGGGTCTCCCCGTGTCAACTGAACTGAGAG 447
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    210 ATCGGTGAGATTGCCACTGGGTCTCCCCGTGTCAACTGAACTGAGAG 269

Qy    448 ATTTCAGTTACCCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 507
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 ATTTCAGTTACCCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCAC 329

Qy    508 AAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACCTGAAAATATATGAACCGCA 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 AAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACCTGAAAATATATGAACCGCA 389

Qy    568 ACTCTAAAAAGATACCAAAACACAAGAGAATAAAATTAGAGTCTTTGGAGAAAATCCCAA 627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 ACTCTAAAAAGATACCAAAACACAAGAGAATAAAATTAGAGTCTTTGGAGAAAATCCCAA 449

Qy    628 GCTGAGTTGAAGAAGGAAGCCAAGAACGAAAGCGAAACGCACTCAAATATGAA 687
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 GCTGAGTTGAAGAAGGAAGCCAAGAACGAAAGCGAAACGCACTCAAATATGAA 509

Qy    688 CACAAAAAATTGAGTTATGTTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 CACAAAAAATTGAGTTATGTTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTC 569

Qy    748 ATTGCAGATGGTTGTCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGGCTTCTGCTTTCTCGTT 807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 ATTGCAGATGGTTGTCAAAGAGGCTCTGCTTGAAGAGAAGAGGCGGCTTCTGCTTTCTCGTT 629

Qy    808 GATAAGGCACTGTGGCCTTTGCAAACCACTACATTACTTACAGTCTGCAGAACTA 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 GATAAGGCACTGTGGCCTTTGCAAACCACTACATTACTTACAGTCTGCAGAACTA 689

Qy    868 CTGAATTCCAAGCTGCCTCGGTGGTCAGGCAGGAGACCTGTTGATGCCATCAAAAGTGCCAGAG 927
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 CTGAATTCCAAGCTGCCTCGGTGGTCAGGCAGGAGACCTGTTGATGCCATCAAAAGTGCCAGAG 749

Qy    928 AAAATCATGAATATGATCGGAGAAAATAAAGACCCCGGTGTCTGGAAACT 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 AAAATCATGAATATGATCGGAGAAAATAAAGACCCCGGTGTCTGGAAACT 809

Qy    988 CCTCAGGGCTTCACCCATGATCGGAGAAGCAATGTGGTTAGGGAAAGATTACGACACCCTT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 CCTCAGGGCTTCACCCATGATCGGAGAAGCAATGTGGTTAGGGAAAGATTACGACACCCTT 869

Qy   1048 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGGAGCATATACCAGTCCCTTG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    870 TCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCAGGAGCATATACCAGTCCCTTG 929

Qy   1108 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCCGGAATTCACAAAAGGGTAAATAAATTCA 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    930 ATCGATATGTTTAATAACCCAGCCACGGCTGCCCCCGGAATTCACAAAAGGGTAAATAAATTCA 989

Qy   1168 ACAGGTACTTCCGAAGATCCCAGTTTACACGGGATCAGTTTGCAACGGGACTGAAC 1227
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990 ACAGGTACTTCCGAAGATCCCAGTTTACACGGGATCAGTTTGCAACGGGACTGAAC 1049

Qy   1228 ATGATGAAGAAGCAGAAAGTGAAGACCATCTTCTCCCACACTGCGGCTCCAACAAGACC 1287
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 ARGATGAAGAAGCAGAAAGTGAAGACCATCTTCTCCCACACTGCGGCTCCAACAAGACC 1109

Qy   1288 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 TTACTCAGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1169

Qy   1348 TGGCTCTATGGAGAACACGAGCGTGTTCCAAGGCGAGGGGTTGGTTCCCCGGAGAAGTGG 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 TGGCTCTATGGAGAACACGAGCGTGTTCCAAGGCGAGGGGTTGGTTCCCCGGAGAAGTGG 1229

Qy   1408 AAGTTGCTCGGAAGAGAAAAATGAGACAGAAGGAGAACAAGCAGTGACACGAACGGTGCCCAGCCCCCCACACCA 1467
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 AAGTTGCTCGGAAGAGAAAAATGAGACAGAAGGAGAACAAGCAGTGACACGAACGGTGCCCAGCCCCCCACACCA 1289

Qy   1468 GTGAGAAGCATCAGCACCCGTTGAACTTGTCTTGAGAATAGCAGTGTTGTTCATCCCCACCC 1527
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 GTGAGAAGCATCAGCACCCGTTGAACTTGTCTTGAGAATAGCAGTGTTGTTCATCCCCACCC 1349

Qy   1528 GACTACTTGGAATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGACAGCAGATTCGGCCAGG 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 GACTACTTGGAATGCTTATCCATGGGGGCAGCTGCCGACAGGAGACAGCAGATTCGGCCAGG 1409

Qy   1588 ACGACATCCACCCTTTAAGGCCCCAGCCAGCCGTTCCAAGCGCCGGCTTCCTAACGATGCC 1647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 ACGACATCCACCCTTTAAGGCCCCAGCCAGCCGTTCCAAGCGCCGGCTTCCTAACGATGCC 1469

Qy   1648 AACGGGGACTTGACAAAGCCGCCTTTTCTCAGGGGGAGAAAAACCCTTTGCCACTGTGAAACTC 1707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 AACGGGGACTTGACAAAGCCGCCTTTTCTCAGGGGGAGAAAAACCCTTTGCCACTGTGAAACTC 1529

Qy   1708 CGCCCGACTGTGACGGAATGAATCGGCTCGGCACCCATCATCGATGAGAGGACAGCAGGA 1767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 CGCCCGACTGTGACGGAATGAATCGGCTCGGCACCCATCATCGATGAGAGGACAGCAGGA 1589

Qy   1768 CTCTCCCCGGGGCCCTCCCGTTCTCCCCTGCCGTTCTCCCCTTGCCGGAATGATGGGCGCCATCCTGTCTCTGCCACGT 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 CGCCCGACTGTGACGGAATGATCGGCTCGGCACCCATCATCGATGAGAGGACAGCAGGA 1649

Qy   1828 GCTGACCGGTCGGAGAAGCTTCAGGTGAGTAGTGTGGGGTTGGGTTGGAGGGCCTAACTCTAAATGGTCGGCGGCATCCGTGTCTCGCCACGT 1887
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 GCTGACCGGTCGGGGAAGCTTCAGGTGAGTAGTGTGGGGTTGGGTTGGGA 1709

Qy   1888 TCATGCTTCTCTGTTTCAGGTAGTGGTTGGGTTGGACAAGGTTTCTGCCTTTAAGATAAAAATGAGT 1947
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1710 TCATGCTTCTCTGTTTCAGGTAGTGGTTGGGTTGGACAAGGTTTTAAGATAAAATGAGT 1769

Qy   1948 AATAGTCTAAATGACCAGCTCAGCCATTTAAAAATTATTTTCTTCTTTTATTTCTGTTTCAAGAAAC 2007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1770 AATAGTCTAAATGACCAGCTCAGCCATTTAAAAATTATTTTCTTCTTTTATTTCTGTTTCAAGAAAC 1829

Qy   2008 AGTAAACTTGGTTTCAATCTTTTA 2030
          |||||||||||||||||||||||||
Db   1830 AGTAAACTTGGTTTCAATCTTTTA 1852
```

RESULT 4

LOCUS   AK027109   1580 bp   mRNA   linear   PRI 12-SEP-2003
DEFINITION   Homo sapiens cDNA: FLJ23456 fis, clone HSI07211, highly similar to
             AF119666 Homo sapiens insulin receptor tyrosine kinase substrate
             mRNA.
ACCESSION   AK027109
VERSION   AK027109.1 GI:10440150
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
             Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
             Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
             Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
             Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE   NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1580)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
             Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
             University of Tokyo, Laboratory of Genome Structure Analysis, Human
             Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
             Japan (E-mail:fldcna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
             Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of

```
FEATURES
source                International Trade and Industry of Japan; cDNA full insert
                      sequencing; Research Association for Biotechnology; cDNA library
                      construction, 5'- & 3'-end one pass sequencing; Department of
                      Virology and Human Genome Center, Institute of Medical Science,
                      University of Tokyo (partly supported by Science and Technology
                      Agency).

                      Location/Qualifiers
                      1. .1580
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /clone="HSI07211"
                        /db_xref="taxon:9606"
                        /clone_lib="HSI"
                        /tissue_type="human small intestine"
                        /note="cloning vector pME18SFL3"

misc_feature          1. .1580
                        /note="highly similar to AF119666 Homo sapiens insulin
                        receptor tyrosine kinase substrate mRNA"

ORIGIN

Query Match    74.4%;   Score 1560;   DB 8;   Length 1580;
Best Local Similarity   99.6%;   Pred. No. 1.4e-260;
Matches 1574;  Conservative  0;  Mismatches  5;  Indels  1;  Gaps  1;

Qy   499 AAATTCCAAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGTGAAATAT 558
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AAATTCCAAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGTGAAATAT 60

Qy   559 ATGAACGCAACTCTAAAAAAGATACCAAACACAGAACACAAAGAAATAAATTAGAGTCTTTGGAG 618
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATGAACGCAACTCTAAAAAAGATACCAAACACAGAACACAAAGAAATAAATTAGAGTCTTTGGAG 120

Qy   619 AAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAGCCAAGGAAACGCACTC 678
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAGCCAAGGAAACGCACTC 180

Qy   679 AAAATATGAACACAAAGAAAAATTGAGTATGTGGAGAACCGTTACTTCGTCAGAGTGAAATC 738
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAAATATGAACACAAAGAAAAATTGAGTATGTGGAGAACCGTTACTTCGTCAGAGTGAAATC 240

Qy   739 CAGAAAATTCATTGCCAGATGGTTGCAAAGAGGCTTGCTTGAAGAGAAGAGGCGCTTCTGC 798
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAAATTCATTGCCAGATGGTTGCAAAGAGGCTTGCTTGAAGAGAAGAGGCGCTTCTGC 300

Qy   799 TTTCTGGTTGATAAGGCACTGTGCCTTTGCAAAACACATACATTATTATACACTTACAGTCT 858
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TTTCTGGTTGATAAGGCACTGTGCCTTTGCAAAACACATACATTATTATACACTTACAGTCT 360

Qy   859 GCAGAACTACTGAATTCCAAGCTTGCCTCGGTGTGGCCAGGAGACCGTGTTGTTGATGCCATCAAA 918
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCAGAACTACTGAATTCCAAGCTTGCCTCGGTGTGGCCAGGAGACCGTGTTGTTGATGCCATCAAA 420

Qy   919 GTGCCAGAGAAAATCATGAATATATGATCGAAGAAATAAAAGACCCAGGCTCTACCCCGTG 978
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTGCCAGAGAAAATCATGAATATATGATCGAAGAAATAAAAGACCCAGGCTCTACCCCGTG 480

Qy   979 TCTCGGAACTCCCTCAGGCTTCACCCATGATCGAGAGAAGAAGCATATTGGTTAGGAAAGATTAC 1038
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TCTCGGAACTCCCTCAGGCTTCACCCATGATCGAGAGAAGAAGCATATTGGTTAGGAAAGATTAC 540

Qy  1039 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCCGCTCCCTTCAGCGAGACATATACC 1098
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCCGCTCCCTTCAGCGAGACATATACC 600

Qy  1099 AGTCCCTTGATCGATAGTTTAATAAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 1158
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGTCCCTTGATCGATAGTTTAATAAACCCAGCCACGGCTGCCCCGAATTCACAAAGGGTA 660

Qy  1159 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACACGGATTTTCGGGTTGCAACG 1218
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACACGGATTTTCGGGTTGCAACG 720

Qy  1219 GGACTGAACATGATGATGAAGAAGCAGAAAGTGAAGACCATCTCTCCGCACACTGCGCGGGCTCC 1278
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GGACTGAACATGATGATGAAGAAGCAGAAAGTGAAGACCATCTCTCCGCACACTGCGCGGGCTCC 780

Qy  1279 AACAAGAACCTTACTCAGCTTTGCACAGGGAGGAGGAGTGTCATCACCGTTGCTGCTTCCCGGACTC 1338
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AACAAGAACCTTACTCAGCTTTGCACAGGGAGGAGGAGTGTCATCACCGTTGCTGCTTCCCGGACTC 840

Qy  1339 AAGGATGGGCTGGGCTCTATGGAGGAACACGACGCAGGGGAGGAGGGTTGGTTCCGTCGG 1398
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AAGGATGGGCTGGGCTCTATGGAGGAACACGACGCAGGGGAGGAGGGTTGGTTCCGTCGG 900

Qy  1399 TCGTACACGGAAGTTGCTGGAACAGAAAAATGAGACAAGAAGCAGAAGCAGAGTGACCGTGCCCAAGC 1458
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCGTACACGGAAGTTGCTGGAACAGAAAAATGAGACAAGAAGCAGAAGCAGAGTGACCGTGCCCAAGC 960

Qy  1459 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTTGTCATC 1518
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTTGTCATC 1020

Qy  1519 CCCCCACCCGACTACTTGGGAATGCTTGTCTCACATGGGGGGCGGCGGCGCGCAGAGAGCAGAT 1578
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCCCCACCCGACTACTTGGGAATGCTTGTCTCACATGGGGGGCGGCGGCGCGCAGAGAGCAGAT 1080

Qy  1579 TCGGCCAGGACGACGACATCCACCCTTTAAGCCCCAGCCGGTCCAAGCCGGAGACCGGGGGCTCT 1638
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCGGCCAGGACGACGACATCCACCCTTTAAGCCCCAGCCGGTCCAAGCCGGAGACCGGGGGCTCT 1140

Qy  1639 AACGGATGCCAACGGGACGTGGAAAGCCGGCCTTTCTCACGTGGGAGAAAAACCCCTTTGCCACT 1698
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 AACGGATGCCAACGGGACGTGGAAAGCCGGCCTTTCTCACGTGGGAGAAAAACCCCTTTGCCACT 1200

Qy  1699 GTGAAACTCCGCCCGACTGTGACGAATGATGCGTCTCGGCGACACCATCATTCGGATAGAGGAC 1758
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGAAACTCCGCCCGACTGTGACGAATGATGCGTCTCGGCGACACCATCATTCGGATAGAGGAC 1260

Qy  1759 AGCCAAGGACTCTCCCGGGGCCCTCTCCGGGTTCTCCCTTGGGGAATGAGTGGGGCGCCATCCTGT 1818
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGCCAAGGACTCTCCCGGGGCCCTCTCCGGGTTCTCCCTTGGGGAATGAGTGGGGCGCCATCCTGT 1320

Qy  1819 CTGCCACGGTGCTGACGGGTGGGGAAGCTTCAGTGGAGAGGGCCTAACTCTAATGTGCCCTGC 1878
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CTGCCACGGTGCTGACGGGTGGGGAAGCTTCAGTGGAGAGGGCCTAACTCTAATGTGCCCTGC 1380

Qy  1879 TTAAGCAAATCATG-CTTCCTCGTTTCACGTAGTAGTTGACAAGTTCTTCCTATTCTG 1937
         ||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db  1381 TTAAGCAAATCATGACTTCTCGTTTCACGTAGTAGTTGACAAGTTCTTGCCTTTAAG 1440

Qy  1938 ATAAATGAGTAAATAGTCTGAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1997
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATAAATGAGTAAATAGTCTGAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTG 1500

Qy  1998 TTCAAGAAACAGTAAAACTTGGTTTCAAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAA 2057
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 TTCAAGAAACAGTAAAACTTGGTTTCAAATCTTTACTGTAAAAAAAAAAAAAAAA 1560

Qy  2058 AAAAAAAAAAAAAAAAAAAA 2077
         |||||||||||||||||||
Db  1561 AAAAAAAAAAAAAAAAAAAA 1580
```

RESULT 5
LOCUS   CQ730103   1491 bp   DNA   linear   PAT 03-FEB-2004
DEFINITION Sequence 16037 from Patent WO02068579.
ACCESSION   CQ730103
VERSION   CQ730103.1 GI:42303049
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1

```
AUTHORS     Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE       kits, such as nucleic acid arrays, comprising a majority of
            humanhexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL     Patent: WO 02068579-A 16037 06-SEP-2002;
            PE Corporation (NY) (US);
FEATURES             Location/Qualifiers
source      1..1491
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN

Query Match          58.9%; Score 1235.4; DB 6; Length 1491;
Best Local Similarity 88.8%; Pred. No. 3e-204;
Matches 1436; Conservative  0; Mismatches  1; Indels 180; Gaps   2;

Qy     61 CGGAGACGCGCGCGGGCGGCGCCGCCACCTCTTCGGGGCCGCCGACCCTTCCAGAGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGAGACGCGCGCGGGCGGCGCCGCCACCTCTTCGGGGCCGCCGACCCTTCCAGAGC 60

Qy    121 CCCTTAGTCGCGCCCCCCCCCCCGCTCCGCCGGAGTCGCCGGCGGGGGGGCCCAGCCG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCCTTAGTCGCGGGGGGCCCTCCCGGCTCCGCCGGAGTCGCCCCACGGACCCCAGCCG 120

Qy    181 CGTCCTCCCCGCGGCCTTCGTCTCGCCCCCCCCCGAGCCCATGTCCCCGGAGGAGGTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGTCCTCCCCGCGGCCTTCGTCTCGCCCCCGCGGCCCCATGTCCCGGAGGAGGTG 180

Qy    241 AACCCGCCTCACGGAGCAGCACCTACCCGGAAAGTGTTAGGGAACAGTTCAATCCTGGG 300
          ||||||||||||||||||||||||||||||||||||||||||
Db    181 AACCCGCTCACGGAGCAGCACCTACCCGGAAAGTGTTATGGAACAGTTCAATCCTGGG 240

Qy    301 AATTTAATAAACCTGGGGAAAAAATTATGAGAAAGCTACTTCCGAAGATCCCAGTTTAC 360
          |||||||||||||||||||||||||||||||||||||||||||||
Db    241 AATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGA 300

Qy    361 AAAGCCTACTACGATGGAGGATCGGTTGAGATTGCCACTGGGTCCCCGTGTCA 420
          ||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAAGCCTACTACGATGGAGGATCGGTTGAGATTGCCACTGGGTCCCCGTGTCA 360

Qy    421 ACTGAACTGGGACAGTGTCCCATAGGATTTCAGTAGCCCAAGATCTCAACGAGAGT 480
          ||||||||||||||||||||||||||||||||||||||||
Db    361 ACTGAACTGGGACAGTGTCTCCATAGGATTTCAGTAGCCCAAGATCTCAACGAGAGT 420

Qy    481 CTTGATGAAAATTTTAAAAAATTCCAAAAGAGATTATCCATGAGCTGGACAAGGACATA 540
          |||||||||||||||||||||||||||||||||||||||||
Db    421 CTTGATGAAAATTTTAAAAAATTCCAAAAGAGATTATCCAAGAGCTGGACAAGGACATA 480

Qy    541 GAACTTGACGTGAAATATATGAAGCAACTCTAAAAAGATACCAAACACAACAAGAAT 600
          ||||||||||||||||||||
Db    481 GAACTTGACGTGAAATATATGAAC-------------------------------- 504

Qy    601 AAATTAGAGTCTTTGGAGAAATCCAAGCTGAGTTGAAGAGATCAGAAGGAAAAGCCAA 660
Db    504 ----------------------------------------------

Qy    661 GGAAGCCGAAACGCACTCAAATATGAACACAAAGAAATTCAGTATGTGAGACCGTTACT 720
          |||||||||||
Db    505 -----------TATGTGGAGACCGTTACT 522

Qy    721 TCTCGTCAGAGTGAAAATCCAGAAATTCATTGCACAGTGGTTGCAAAGAGAGGCTCTCTTGTTGAA 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 TCTCGTCAGAGTGAAAATCCAGAAATTCATTGCACAGTGGTTGCAAAGAGAGGCTCTCTTGTTGAA 582

Qy    781 GAGAAGAGGCGCGCTTCTGCTTCTTGCTTTCGGTTGTGATAAGGCACTCGAAAGGCTTTTCAAACCACACATAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 GAGAAGAGGCGCGCTTCTGCTTCTTGCTTTCGGTTGTGATAAGGCACTCGAAAGGCTTTTCAAACCACACATAC 642

Qy    841 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCTTCGGTTGGCAGGAGACC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    643 TATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCTTCGGTTGGCAGGAGACC 702
```

RESULT 6
BC015459

```
LOCUS       BC015459                3040 bp    mRNA    linear   ROD 01-MAR-2005
DEFINITION  Mus musculus BAII-associated protein 2-like 1, mRNA (cDNA clone
            MGC:8269 IMAGE:3592632), complete cds.
ACCESSION   BC015459
VERSION     BC015459.1  GI:15930030
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3040)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
```

```
Qy    901 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATTGAATATGATGAAGAAAATAAGACC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    703 TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATTGAATATGATGAAGAAAATAAGACC 762

Qy    961 CCAGCCTCTACCCCCGTGTCTTGGAACTCCTTCAGGCGCTTCACCCATGATCGAGAAGCAAT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    763 CCAGCCTCTACCCCCGTGTCTTGGAACTCCTTCAGGCGCTTCACCCATGATCGAGAAGCAAT 822

Qy   1021 GTGGTTAGGAGAAGATTACGACACACCCTTTCTAAATGCTCACCAAAAGATGTCCCCGCCTCTT 1080
          ||
Db    823 GT---------------------------------GATGCCCCCGCCTCCT 840

Qy   1081 TCAGGCAGAGCAATATACCAGTGCCCTTGATCGATCGAGCTGGTTTAATAACCAGCCACACGGCGGCTGCC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCAGGCAGAGCAATATACCAGTGCCCTTGATCGATCGGTTTAATAACCAGCCACACGGCGGCTGCC 900

Qy   1141 CCGAATTCACAAAGGGGTAAATAAATTCAACAGGTGACTTCCGAAGTACTTCCGAAGTTTACAGGCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCGAATTCACAAAGGGTAAATAATTCAACAGTTCCCTTGATCGATCGGTTTAATAACCAGCCACACGGCA 960

Qy   1201 TCAGTTTCGGTTGCCAAGGGGACTGAACATGAAGAAGAGCAGAAAGTGAAGACCATCTTC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCAGTTTCGGTTGCCAAGGGGACTGAACATGAAGAAGAGCAGAAAGTGAAGACCATCTTC 1020

Qy   1261 CCGCACACTGCGGGCTCCAACAAGAACCTTACTCAGCTTTGCACAGGGGGAGATGTCATCACG 1320
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCACACTGCGGGCTCCAACAAGAACCTTACTCAGCTTTGCACAGGGGGAGATGTCATCACG 1080

Qy   1321 CTGCTCATCCCCGGAGGAAGAAGGATGGCCTGGCTTGGCTTGGCTCCTATGGAGACACGACGTGTCCAAGCG 1380
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTGCTCATCCCCGGAGGAAGAAGGATGGCCTGGCTTGGCTTGGCTCCTATGGAGACACGACGTGTCCAAGCG 1140

Qy   1381 AGGGGGTTGGGTTCCCGTTGCGTTCGTTCGTACACCAGAAGTTGCTGGAAGAAAATGACACAGAAGCAGTG 1440
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGGGGGTTGGGTTCCCGTTGCGTTCGTTCGTACACCAGAAGTTGCTGGAAGAAAATGACACAGAAGCAGTG 1200

Qy   1441 ACCGTTGCCACCGCCCAAGCCCACCACCACCAGTGAGAAGCAGAAAGTGGAAGTTGTCTGAG 1500
          |||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACCGTTGCCACCGCCCAAGCCCACCACCACCAGTGAGAAGCAGAAAGTGGAAGTTGTCTGAG 1260

Qy   1501 AATAGCACGTTGTCATCCCCGCCCAGGACGACATCCACCTTTGTTGTCCATGGGGGGGCAGCT 1560
          ||||||||||||||||||||||||||||||||||||||||||
Db   1261 AATAGCACGTTGTCATCCCCGCCCAGGACGACATCCACCTTTGTTGTCCATGGGGGGGCAGCT 1320

Qy   1561 GCCGACAGGAGAGCACGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCACCGTCCAAG 1620
          |||||||||||||||||||||||||||||||||||||||||
Db   1321 GCCGACAGGAGAGCACGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCACCGTCCAAG 1380

Qy   1621 CCCGAGACGCCGGGCTCCTACGATGCCAACGGGACTGCCAAAGCCGCCGCCTTTTCTCAGC 1677
          |||||||||||||||||||||||||||||||||||||||||
Db   1381 CCCGAGACGCCGGGCTCCTACGATGCCAACGGGACTGCCAAAGCCGCCGCCTTTTCTCAGC 1437
```

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE   Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences

JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED   12477932

REFERENCE 2 (bases 1 to 3040)

AUTHORS   Director MGC Project.

TITLE   Direct Submission

JOURNAL   Submitted (01-OCT-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA

REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT   Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site: http://www-shgc.stanford.edu
          Contact: (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 11 Row: b Column: 23
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 27229018.

FEATURES   Location/Qualifiers
     source   1..3040
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /clone="MGC:8269 IMAGE:3592632"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene   1..3040
                     /gene="Baiap2l1"
                     /db_xref="GeneID:66898"
                     /db_xref="MGI:1914148"
     CDS   50..1594
                     /gene="Baiap2l1"
                     /codon_start=1
                     /product="BAI1-associated protein 2-like 1"
                     /protein_id="AAH15459.1"
                     /db_xref="GI:15930031"
                     /db_xref="GeneID:66898"
                     /db_xref="MGI:1914148"
                     /translation="MSRGPEEVNRLTENTYRNVMEQPNPGLRNLINLGKNYEKAVNAM
                     ILAGKAYVDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNETLDENFKKFHKDIIH
                     ELEKKTELDVKYMNATLKRYQAEHRNKLDSLEKSQAELKKIRKSQGRNALKYEHKE
                     IEYVETVTSRQSEIQKRIADGGEKEALBEKRPCFLVDKHGSFASHIYYHMQSAELL
                     NSKLPRWQETCCDATKVPEKIMNMIEIKTPISTPVSGTPQPSPMIERSKMIGKDYDT
                     LSKYSPKMPPAPSVKAYTSPLIDMFNPATAAQSSEKTNNSTANTGEDPSLQRSVSVA
                     TGLNMMKKQKVKTIPPHPAGNNKTLLLSPAQGDVLTLLIPEEKDGWLYGEHDTKARGW
                     FPSSYTKLLEENEAMSVPTPSPAPVRSISTVDLTEKSSVVIPPDYLECLSMGATSDK
                     KAGAPKVPSASTPRAPVSRPDATSTSPSDANGTRAKPPFLSGENPPATVKLRPTVTNDR
                     SAPIIR"

ORIGIN
Query Match   55.5%; Score 1163; DB 9; Length 3040;
Best Local Similarity   78.5%; Pred. No. 1.2e-191;
Matches 1501; Conservative 0; Mismatches 380; Indels 30; Gaps 8;

Qy   168 CGAGGGCCCAGCGCGCGTCCTCCCCGCGCGTTGCTCGCCGCGCCCAGCCATGTCCCGGGG 227
          |||||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1 CGAGGGCCCCGCGCGCGTCCTCCCCGCGCGTTGCTCGCCGCGCCCAGCCATGTCCCGGGG   60

Qy   228 GCCCGAGGAGGTGAACCGGCTCACCGAGAACACCTACCGGAATGTTATGGAACAGTTCAA 287
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCCGGAGGAGGTGAACCGGCTCACCGAGAACACGTACCGGAATGTGATGGAACAGTTCAA 120

Qy   288 TCCTGGGCTGCGAAATTTAATAAACCTGGGAAAAATTATGAGAAAGCTGTAAACCTAT   347
          |||||||||||| |||||||||||||||||||||||| |||||||||||||| |||||
Db   121 CCCTGGACTACGGACTTAATAAACCTAGGAAAAATTATGAGAAAGCGGTGAATGCTAT   180

Qy   348 GATCCTTGGCAGGAAAAAGCCTACTACGATGAGTGGCCAAGATCGGTGGAGATGCCACTGG 407
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   181 GATCCTGGCAGGAAAAAGCCTACTACGACGAGTGGCCAAGATCGGTGGAGATGCCACTGG 240

Qy   408 GTCCCCGGTGTCAACTGGACACAGTGTCCTCATAGAGATTTCAGTGTACCCACAAGAGA 467
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTCTCCGGTGTCAACAGAGCTAGGCCATGTCCTCATAGAGATTTCCAGCCACCACAAGAA 300

Qy   468 ACTCAACCGAGAGTCTTGATGAAAATTTTAAAAAATTCCCACAAGAGAGATTATCCATGAGCT 527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACTCAATGACGACTTCTTGATGAAAATTTCAAAAAATTCCATAAGGATATTATCCATGACTT 360

Qy   528 GGAGAAGAGAGATAGAACTTGACGTGAAAATATATGAAAGGCAAACTCTAAAAAGATACCAAAC 587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGAGAAGAAGACAGAACTTGATGTGAAATATATGAACGCTACCCTACTCTTAAAAGATATCAAGC 420

Qy   588 AGAACACAAGAGAATTAAATTGAGTCTTTGGAGAAATCCAGAAAATCCCAAGCTGAGTTGAAGAAGATCAG 647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGAGCACAGGAATAAATTAGATTCTTTGGAGAAATCCAGGGCGGAGCTGAAGAAGATCAG 480

Qy   648 AAGGAAAAAGCCAAGGAAGCCCGAAACGCACTCAAATATGGAACACAAAGAAAATTGAGTATGT 707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGAAAAAGTCAAGGTGACGAAAACGCACTCAAATATGAACATAAAGAAATTGAGTATGT 540

Qy   708 GGAGACGTTACTTCGTCAGAGTGAAATCCAGAAAATCATTGCAGATGGTGGTTGCAAGA 767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GGAGACGGTTGACGTCTCGCCAGAGTGAGATCCAGAAGCGAATCATTGCCGATGGGGTTGCAAAGA 600

Qy   768 GGCTCTGCTTGAGCAGAGACGGCGCCGCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGCCCTGCTTGAAGCAGAAAAAGGCGCCGCTTCTGCTTTCTGGTGGACAAGCACTGCAGCTTTGC 660

Qy   828 AAACCACATACATTATTATTCACATTACACTTACAGTCTGCAGAACTACTGAATTCAAAGCTGCTCG 887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CAGTCACATACACTTATTACCACATGCAGTCTGCAGAGTTACTTAATTCCAAGCTGCCCCG 720

Qy   888 GTGGCAAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATATGATCGA 947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GTGGCAGGAGAAACCTGTTGTGATCCACAAAGTCCCAGAGAAAATCATGAACATGATAGA 780

Qy   948 AGAAATAAACAGACCCAGCCTCTACCCCCGTGTCTCGGAACTGCTCAGGGTTCACCCATGAT 1007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGAGATAAAGACGCCCATCTCACCCGTGTCCGGGAACCCCACACCCTCAGCCTTCACCATGAT 840

Qy   1008 CGAGAGGAGCAAATGTGGTTAGGGAAAGGATTACGACACCCTTTCTAAATGTCACCAAAGAT 1067
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGAGAGGAAGCAAGATGTGGGAGAAGATTATGACACCCTTTCTAAATGCTACCAACAAAGAT 900

Qy   1068 GCCCCCGCTCCTTCAGGCAGAGCAATATCCAGTCCCTTGATCGATATGTTTAATACCC 1127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GCCCCCTGCTCCTTCAGTCAAAGCCTACACCAGTCCCTTGATTGACATGTTCAATAACCC 960

Qy   1128 AGCCACGGGCTGCCCCGGAATTCACAAAGGGTAAATAATTCAACAG---GTACTTCCGAAGA 1184
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   961 AGCCACGGGCTGCCCCGGAATTCACAAAGGGTAAATAATTCAACAG   GTACTTCCGAAGA 1184

RESULT 7
BC089216

```
LOCUS       BC089216                2141 bp    mRNA    linear   ROD 30-MAR-2005
DEFINITION  Rattus norvegicus similar to RIKEN cDNA 130006M19 (predicted),
            mRNA (cDNA clone IMAGE:7303117), partial cds.
ACCESSION   BC089216
VERSION     BC089216.1  GI:58402628
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 2141)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2141)
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
            College of Wisconsin
            cDNA Library Preparation: Express Genomics
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 202 Row: 1 Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis. Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1..2141
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7303117"
                     /tissue_type="Ovary, rat"
                     /clone_lib="NIH_MGC_253"
                     /lab_host="DH10B"
                     /note="Vector: pExpress1"
     gene            <1..2141
                     /gene="RGD1308452_predicted"
```

```
Db    961 AGCTACAGCTGCACAGAGTTCAGAGAAAAAAACAAACTCCACAGCAAACACTCGGGAAGA 1020
              |||   |||       ||||
Qy   1185 TCCCAGTTTACAGCCGATCACTTTCGGTTGCAACGGACTGAACATGATGAAGAAGCAGAA 1244
          ||||||||||||||||||||||||
Db   1021 TCCCAGTCTACAGCGGATCCGGTTTCTGTTGCAACTGGGCTGAACATGATGAAGAAGCAGA 1080

Qy   1245 AGTGAAGACCATCTTCCTCCCACACTGCCGCGGCTCAAACAAGACCTTACTCAGCTTTGCACA 1304
          ||||||||||||||||||   ||||||
Db   1081 GGTGAAGACCATCTTCCTCACACGGCTGCCAACAATAAGACCCTACTGAGCTTTGCCCA 1140

Qy   1305 GGGGAGAGTGTCATCACGGCTCCACTGCCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAACA 1364
          ||||   |||||||||||   ||||
Db   1141 GGGGGGATGTCCTCACACTGCTCATCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGCA 1200

Qy   1365 CGACGTGTCCAAGGCGGAGGGGGTTGGTTCCGGTGACGAGAGTTGCTGGAAGAAA 1424
          |||  ||||||||   |||||
Db   1201 TGACACCACCAAAGCGAGGGCTTGGTTCCCATCGTATATACAAAAGGTTGCTGGAAGAAAA 1260

Qy   1425 TGAGACAGAAGCAGTGACCGTGCCACACCCCACAGCCCAGCAGTGGAAGCATCAGCAC 1484
          |     ||||||   |||||
Db   1261 T-----GAAGCAAATGTCTGTCCTACACAAGCCCTGCACCCGTGAGAAGCATCAGCAC 1314

Qy   1485 CCGTGAACTTGTCTGAGAATAGCAGTGTTGTATCATCCCCCAGACTACTTGGAAATGCTT 1544
          |||||   |||||||
Db   1315 GGTGGGACCTGACTGAGAAGAGCAGCGGTTGTCATCCCGACTACTTGGAGTGCCT 1374

Qy   1545 GTCCCATCGGGGCAGCTGCCGACAGGAGGAGCAGATTCGCCCAGG------ACCGACATCCAC 1598
          ||||   |||||
Db   1375 GTCCATGGGAGCCCCACCTCAGACAAGGAAAAGGCCTCCCAAGGTGCCTTCTGCTTCTAC 1434

Qy   1599 CTTTAAGGCCCCAGGGTTCCAAGCCCGA-------GACCGCGGGCTCCTAACGAGCCCAACGG 1652
          |||     |||||
Db   1435 CTTCAGAGCGCGCCAGTGTCACGACCAGCAGCAGACGCCCACATCCCACGGATGCCAACGG 1494

Qy   1653 GACTGCAAAGCGCGCCCTTTTCTCAGCGGAGAAAACCCCTTGCCACTGTGAAAACTCGCCC 1712
          |||  ||||||||   |||||
Db   1495 GACCGCAAAGCGCCCATTCCTCAGTGAGAAAAACCCTTTGCCACCGTGAAAACTCCGGCC 1554

Qy   1713 GACTGTGACGAATGATCGCTCGGGCACCCATCATTCGGATG-AGAGGACAGCCAAGGAGACTCT 1771
          |||  ||||||   |||||
Db   1555 AACAGTGACAAATGACCGATCAGCACCGATGAAGACGAAGCAGGGACGACAGTGCCA 1614

Qy   1772 CCCGGGCCTCTCCCGGTTCTCCCCTTGCGGGAATGAGTGGGCGCAATCCTGTCTCGCCACGTGCTG 1831
          ||||   |||||
Db   1615 GTCCCCTCACTCGAGTTCTCACCCACAGGGGCCGAGTAGGTATA-CAATTGCTGGACAGCAGGC 1673

Qy   1832 ACGGTCGGGGAGCTTCAGTGGAGAGGCCTACTCTTAAGTGTCCCCTGCTTAAGCAAATCAT 1891
          ||||   |||||
Db   1674 ACTGTCAAAAGATTGCAGAGTGCCTGTGTCTCAATGAAGCCCCACTTCAGCAAATCAATGTA 1733

Qy   1892 GCTTCTCTGTTTCACGTTAGTTGGGTTGACAGTTTCTGCCTTTAAGATAAATGAGTAATA 1951
          ||||   |||||
Db   1734 CTTTCTCAGATTGACAGAGGTGGGGTTTACACATTTCTCTTTAAGGTAAAATGAG---TA 1790

Qy   1952 GTCTAAATGACCAGCCATTTAAAAATATTTTTCTTCATTTACTTCAAGAAAAACAGATA 2011
          ||||   |||||
Db   1791 GCTTAGTGAATAAAATCAATCAATTGAAATATTTTCTT----TTCTATTCAAAAAGCAGTA 1846

Qy   2012 AACTTCGGTTTCAAATCTTTAAAAAAAAATATTTTTAAAAAAAAAAAAAAAAAAAAAAAA 2062
          ||||   |||||
Db   1847 AATTGGCTCAAATTTTTATTTTTAAGGAAAAAATAGCCAGAAATA 1897
```

```
CDS             /db_xref="GeneID:304282"
                /db_xref="RGD:1308452"
                /gene="RGD1308452"
                <1..1620
                /codon_start=1
                /product="RGD1308452_predicted"
                /protein_id="AAH89216.1"
                /db_xref="GI:58402629"
                /db_xref="GeneID:304282"
                /db_xref="RGD:1308452"
                /translation="MFPSADSCGGSRLPRAQSQALGSPAARSPEATRPRRVLPRSPS
                RRLHVSGAGGGEPAHGEHVPAMILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEIS
                STHKKLNETLDENPKFHKEIIHELEKKTELDVKVMNATLKRYQAEHRNKLDSLEKSQ
                AELKKIRRKSQGGRNALKEHKEIIEYVETVTSRQSEIQKFIADGCKEALLEEKRRFCF
                LVDKHCSFASHIHRYHLQSAELLNSKLPRWQETCCDATKVPEKIKNMIEIKTPISTP
                VSGTPOPSPMTERSRKMIGKDYDTLSKYSPKMPPAPSVKAYTSPLIDMFRNPATAQSA
                EKTNNSTANYGDDPSLQRSVSVATGLNMMKKQKVKTIFPHTAGNNKTLLSFAQGDVLT
                LLIPEEKDVGHWLYGEHDITKVRGWPFSSYTKLLEENMKEAMSVPTPSSAPVRSISTVDL
                TEKSSVVIPPDYLECLSMGATSDKRADAAKIPSTSTFKAPVPRPDATSTSPSDSNGT
                AKPPFLSGENPFATVKLRPTVTNDRSAPIIR"

ORIGIN

Query Match     49.8%;  Score 1042.8;  DB 9;  Length 2141;
Best Local Similarity   75.8%;
Matches 1466;  Conservative 0;  Mismatches 362;  Indels 106;  Gaps 10;
                Pred. No. 8.3e-171;
```

```
Qy     75  GGGCACTCTCGGGGCCGAGCGGCTGCGGCGGGGCGGGGCGGGGCGGCCCCTCCAGAGCCCCTTAGTCGCGGCC   134
           ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db      2  GGTTCTCTCCGGGCTGACAGCTGTGTGGGCGGACGCGACTCCCGAGCGCCCGCTCCAGTCGCGAAG          61

Qy    135  CC---GGCCCTCCCGGCGAGTCCGGCGGCGGGGCCCACGAGAGCGGGGGCGGGGGCCGCTCCTCCCGC        191
           |||    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62  CCCTAGGTTCTCCATAGGACCGCCCCGCGGGGCCCAGTCCCGGCGGCGCCCACCAGGTCCTCCCGC         121

Qy    192  GCTTGCTCTCGCCCCGCGCGCCGCCATGTCCCCGGGGGGCCCGGCGCGCGGGGTGAACCCGGCTCAC        251
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122  GCTTCGCCCTCCCGGAGGCTGCA--CCATGTCTCGGGGGCCCGGAGCGGGTGAACCCGGCTCAC         180

Qy    252  GGAGAGCACCTACCCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAA         311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  GGAGAACACGTACCG------------------------------------------ 195

Qy    312  CCTGGGGGAAAAATTATGAGGAAGCTGTAAAAGCTTATGATCCTGGCAGGAAAAAGCCTACTA       371
           |||||||||||||||||||||
Db    196  -------------------------GCTATGATCCTGGCAGGAAAAAGCCTACTA         224

Qy    372  CGATGGAGTGGCCAAAGATCGGTGAGATTGGCCACTGGGTCCCCGTGTCAACTGAACTGGGG       431
           |||| || |||||  |||||||  ||||||||||||||||||||||||||||||||||
Db    225  CGACGGAGTGGCCAAAGATTGGGGAGAAATTCGGCCACTGGGTCTCTGTGTCAACAGAACTAGG         284

Qy    432  ACATGTCTTCATAGAGATTTCAGTAACTCAACAGAGAGTCTTGATGAGAAA         491
           || | |||||||||||||||||||||||||||||||||||||
Db    285  CCAAGTTCTTCCATAGACATTTCCAGCACCCAACAAGAAACTCAACGGAGACTCTTCGATGAGAA         344

Qy    492  TTTTAAAAATTCCACAAAGAGTATATCCATGAGTAGAAGATAGAACTTGACGT         551
           |||||||||||||||||||| || |||||| ||||||||||
Db    345  TTTCAAAAATTCCATAAGGAGATTAATCCATGAGTTGGAGAAGACAGAACTTGATGT         404

Qy    552  GAAATATATGAAACGCAACTCTAAAAAGAGATACCAAACAGAACACACAAGAAATAAATTAGAGTC       611
           || ||| ||||||  ||| ||||||||| ||||||||
Db    405  AAAGTACATGAAACGCTACTCTTAAAAGAGATTATCAAGGCAGAACTGGGAACTGGACAGAATTC         464

Qy    612  TTTTGGAGAAATCCCAAGCTGAGTTGAAGAGATCAGAAGAGCAAAAGCCCAAGGGAAGCCGAA       671
           | ||||||||||||| ||||||  ||||||||||||||| ||||||
Db    465  TTTTGGAGAAATCTCAGGCGGAGTTGAAGAGATCAGAGAGAAAAGTCAAGTGGACGAAA         524

Qy    732  TGAAATCCAGAAAAATTCATTGCAGAATGGTGTTGCAAAGAGGCTCTTGCTTGAAGGAGAAGAGGCG   791
           |||||| |||||||||||||| |||||||||||||||||||||| |||||||||||||||
Db    585  TGAGATTCCAGAGAAGTTCATCGGCAGATGGCCTGCAAAGAGGAGCCTGCTTGAGGAAGAAGAGGCG         644
```

```
Qy    792  CTTCTGCGTTTCTGGTTGATAAGCACACTGTGGCTTTGCAAAACCACATACATTATTCATCTT        851
           ||||| ||||||| |||||||| ||| |||||||| ||||| |||||||||||||||||| ||
Db    645  CTTTCTGCTTTCTGGTTGGATGGAAGCACTGCAGTTTGCCAGTGACATACACCGCTACACT        704

Qy    852  ACAGTCTGCAGAACTACTGAATTCCAAGCTTCCAAAGCTGCCTCGGTGCGGAGGAGGACCTTGTTGGTGGAA        911
           |||||||||||||||||||||||||||| |||||| |||||||||||||||||||||||||||
Db    705  GCAGTCTGCAGAATTACTTAACTCCAAAGCTGCCTCGGTGCAGAGAGGAGGAAGACCTGTTGGTGAGC         764

Qy    912  CATCAAAGTGCCAGAGAAAATCATGAATAATGATCGAGAGAAAATAAAGAGACCCAGCCTCTAC        971
           |||||||||||||||||||| ||||||||||| ||||||||| |||||||||||||
Db    765  CACGAAAGTCCCAGAGAAAATCATGAACATGATAGAGATAAAGAGGACCCCATCCTAC         824

Qy    972  CCCGGTGTCTGGAGAACTCCTCAGGCTTCACCCATGAGTGGAAGCCAATGTGGTTAGGA       1031
           ||||||||||||||||||| |||||| ||||||||||||||||||||||||||
Db    825  CCCAGTGTCCGGAGACGGCCACAGCCATCGGCCATGCACTGATGAGAAGCCAAGATGATTGGGAA         884

Qy   1032  AGATTACGGACACCCTTTCTAAAATGCTCACCAAAGAGGCCCCCCCCCCTTCCTTCAGGCAGAGC       1091
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    885  AGACTACGGACACCCTTTCTAAAATACTCACCGAAGAGATGCCCCCCGCCTCCTTCAGTCAAGGC         944

Qy   1092  ATATACCAGTCCCTTGATCCGATATATTTTAATAACCCAGCCACCGGCGTGCCCGGAATTCACA      1151
           ||||||||||||||||||||||||||||||||||
Db    945  CTACACCAGTCCTCTGATTGAATGTCTAAACAACAACCCTGCCACGGGCGACAGGAGCGCAGA         1004

Qy   1152  AAGGGTAAAATAATTATTCACACAG--GTACTTCCGAAGATCCCAGTTTACAGCGGGATCAGTTTC        1208
           ||||||||||||||||||||||||  |||||||||||||||| ||||||||
Db   1005  GAAAACCAACAACTCCACAGCAGAAACACTGGGGACGACTCCCAGTTTACAGCGGTCTC         1064

Qy   1209  GGTTGCAACGGGGACTGAACATGATGAGAAAGCACAGAAAGTGAAGAACCATCTTCCCTCACAC       1268
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   1065  TGTTGCAACTGGGGCTGAACATGATGAAGAAGCAGAAAGTGAAGAACCATCTTCCTCACAC         1124

Qy   1269  TGCCGGGCTCCAACAAGACTTACTCACGCTTTGCACAGGGAGATGTCATCACGGCTGCTCAT       1328
           ||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db   1125  CGCTGGGCAACAAATAAGACCTTACTGATGGATACAGGGAGAGTGACATCATCAGGCTGCTCAT         1184

Qy   1329  CCCCGAGGAGGAAGGATGGGCTGGCTGTTCCTCTATGGAGAACACAGCAGTGTCCCAAGGGGAGGGGGTG      1388
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1185  TCCCGAGGAGAAAGGACGGCCTGGCTTTACGGGAGGAGCATGACACCAAAGTGAGGGGCTG         1244

Qy   1389  GTTCCCGTGGTCGTGTACCGGAGGTTGCTGGAAGAAAATGAGACGCAGGAGCAGAGTGGACCGGTGCC      1448
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1245  GTTCCCATCATCAATACACAAAAGTTGCTGGAAGAAATGAATGAAGCAATGTCTGTACC         1304

Qy   1449  CACCGCCAAGCCCCACACAGTGAGAAGCAATCAGCACCGTGAACTTGTCTGAGAATAGCAG       1508
           ||||||||||||||||||||||||||||||||||||||||||
Db   1305  AACACCAAAGTCTCGACCACCCAGTGCAGTGAGAAGCATCAGCACTGTGGACCTGACAGAGCAG         1364

Qy   1509  TGTTGTCATCCCCCACCGGACTCACTTGTGAAATGCTTGTTCATGGGGCCGCAGACAGCAGAA       1568
           |||||||||| |||||||||||||||| ||||||||||||||| ||||||||
Db   1365  CGTTGTCATCCTCCCTCCTGACTACCTGGAGTGCCTGTCCATGGGAGCCACCTCCGACAAG         1424

Qy   1569  GAGAGCAGATTCGGGCCCAGGA------CGACATCCACCCTTTAAGCCCCCACAGTCCAAGCC       1622
           ||||||||||||||| |||||      ||||||||||||||| |||||||||||
Db   1425  GAGAGCAGAGCAGACGGTCGCCAAGATACCTTCTACTTCTCAAAAGCCTTCAGCC         1484

Qy   1623  CGAGACCGGCG----------CTCCTACGGATGCCAAACGGGGACTGCAAACGCGGCCTTTTTCTCAG       1676
           |||||||||          ||||||||| |||||| ||||||||
Db   1485  AGATGCCACGTTCCACAATCTCCCAGTGACTCAAAATGGGAACGCCAATTCCTCAG         1544

Qy   1677  CGGAGAGAAAACCCCTTTGCCACTGTGTAAAACTCCGCGACTGTGACGGAAATGCCGCTGCGGC       1736
           |||||||||||||||||||||| |||||||||||||||||||||| |||||||
Db   1545  TGGAGAAAAACCCCTTTGCCACCGTGAAACTCCGGCCCACTGTGACCAATGACCGGTCGGC         1604

Qy   1737  ACCCATCATTCGATGAGGAGGACAGCCAAGGACTCTCCGGCGCCTCCGGTTCTCCCTTG       1796
           ||||||||||| ||||||||||  || |||||||||||||
Db   1605  ACCGATCATCCGGTGAAAATCCAG---AGGAACCTGCCGGAGTGGGTTC---TCTCACCTG         1658

Qy   1797  CGGACTCAAATAATGAACACAAAGAAAATCGAGTTACTGTGGAGACAGTTACCCTCCAGTGAGAG       1856
           ||||||||||||||||| |||||||||||||||
Db   1659  CAGGGGGGATAGGTCTTCACTGCTGGACAAGTGCAGGCCACTGACAAAGCCCTGCAGTGAGAGGGTG         1718

Qy    585  TGAGATTCCAGAGAAGTTCATCGGCAGATGGCCTGCAAAGAGGAGCCTGCTTGAGGAAGAAGAGGCG         644
```

RESULT 8

Qy	1857	GCCTAAACTCTAATGTCGGCCTGCTTAAGCAAATCA--TGCTTCTCTGTTTTCACGTAGTTGG	1914
Db	1719	CCTGGATCTCAATGAACCCCGACTTGAGCAAATCAATGAGTATATTTCTGTTTGATATAGAGCTGG	1778
Qy	1915	GTTGACAGAGTTTCTGCCCTTTAAGATAATAAATGAGTAATAGTCTAATGACCAGCTCAGCCATT	1974
Db	1779	GCTTACACAACAATTTTGTCTTTAAGGTAAAATGAG---TAGCTTAGTGCATAAAATGAATCGTT	1835
Qy	1975	TAAAATATTTTCTT	1988
Db	1836	TGAAGTGTTTTCTT	1849

LOCUS   AK027142   2521 bp   mRNA   linear   PRI 12-SEP-2003
DEFINITION   Homo sapiens cDNA: FLJ23489 fis, clone LNG00746.
ACCESSION   AK027142
VERSION   AK027142.1 GI:10440194
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
                Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
                Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
                Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
                Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE   NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2521)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
                Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
                University of Tokyo, Laboratory of Genome Structure Analysis, Human
                Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
                Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
                Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
                International Trade and Industry of Japan; cDNA full insert
                sequencing; Research Association for Biotechnology; cDNA library
                construction, 5'- & 3'-end one pass sequencing; Department of
                Virology and Human Genome Center, Institute of Medical Science,
                University of Tokyo (partly supported by Science and Technology
                Agency).
FEATURES   Location/Qualifiers
     source   1..2521
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="LNG00746"
                     /tissue_type="human lung"
                     /clone_lib="LNG"
                     /note="cloning vector pME18SFl3"
     CDS   65..589
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="BAB15671.1"
                     /db_xref="GI:10440195"
                     /translation="MMKKQKVKTIPPHTAGSNKTLLSFAQGDVITLLIPEEKDGWLYG
                     EHDVSKARGWFPPSSYTKLLEENETEAVTVPTPSPTPVRSISTVNLSENSSVVIPPDY
                     LECLSMGRAADRRADSVRTTFKVPVFKPEIAAPNDANGTAKPFLSGENPFATVKL
                     RPTVNDRSAPIIR"

RESULT 9

Qy	1164	TTCAACAGGGTACTTCCGAAGATCCCAGAGTTTACAGCAGGATCAGTTTCGGTTGCAAACGGGGACT	1223
Db	1	TTCAACAGGGTACTTCCGAAGATCCCAGAGTTTACAGCAGGATCAGTTTCGGTTGCAAACGGGACT	60
Qy	1224	GAACATGATGAAGAAGCAGAAAGTGAAGAACCATCTTCCCGACACACTCGCGGCGCTCCAACAA	1283
Db	61	GAACATGATGAAGAAGCAGAAAGTGAAGAACCATCTTCCCGACACACTGCGGGGCTCCAACAA	120
Qy	1284	GACCTTACTCAGCTTTTGCACAGGGAGATGATGTCATCACGCTCTCATCCCGAGGAGGAGGA	1343
Db	121	GACCTTACTCAGCTTTGCACAGGGAGATGTCATCACGCTCTCATCCCGAGGAGGA	180
Qy	1344	TGGCTCGGCTTCTCATGTGAGGAACACGACGGTGTCCAAGGGCGAGGGGGTTGGTTCCCGTCGTA	1403
Db	181	TGGCTCGGCTACTTCCGGAAGACACGACGGTGTCCAAGGGCGAGGGGGTTGGTTCCCGTCGTA	240
Qy	1404	CACGGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCAGCCCAAGCCCAC	1463
Db	241	CACGGAAGTTGCTGGAAGAAAATGAGACAGAAGCAGTGACCGTGCCAGCCCAAGCCCAC	300
Qy	1464	ACCAGTGAGAAGCATCAGCACCCGTGAAACTTGTCTGAGAATAGCAGTGTGTTGTCATCCCC	1523
Db	301	ACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTCTGAGAATAGCAGTGTTGTCATCCCC	360
Qy	1524	ACCCCGACTACTTGGAATGCTTGTCCATGGGGGAGGGGGGAGGGGGGAGGGGGGGAGATTCGGC	1583
Db	361	ACCCCGACTACTTGGAATGCTTGTCCATGGGGGAGGGGGGGAGGGGGGGAGATTCGGT	420
Qy	1584	CAGGACGACAATCCACCTTTAAGGCCCCCAGGGTCCAAGCCCGAGACCAGTGACGGGACGGGA	1643
Db	421	CAGGACGACATTCACCTTTAAGGTCCCAGTGTTCAAGCCCGAGACGGGA	480
Qy	1644	TGCCAACGGGGACTGCAAAACCCGCCTTTTCTCAGCGCGAGAAAACCCCTTTGCCCACTGTGCC	1703
Db	481	TGCCAACGGGGACTGCAAAGCCGCCTTTTCTCAGCGCGAGAAAACCCCTTTGCCCACTGTGAA	540
Qy	1704	ACTCCGCCCGACTGTGACCGAATGATCGGCTCGGCGCCATCAGTCGGATGGGATGGGAGGCA	1763
Db	541	ACTCCGCCCGACTGTGACCGAATGATCGGCTCGGCGCCACCCATCAGTCGGATGAGGAGCAGCA	600
Qy	1764	AGGACTCTCCCGGGCGCCCTCCGGGTTGTTCGCGGAATGATGATGGGCGCATCCGTGTCGCC	1823
Db	601	AGGACTCTCCCGGGCGCCCTCCGGTTCTCCGGTTCTCCGGTTCTCGCGCCATCCGTCTCGCC	660
Qy	1824	ACGTGCTGACGGGTCGGGGAAGCTTCAGTGGGAGAGGCCTACTCTAATGTGGGCGCGGTTAAG	1883
Db	661	ACGTGCTGACGGGTCGGGGAAGCTTCAGTGGGAGAGGCCTACTCTAATGTCGCCGTTAAG	720
Qy	1884	CAAATCATGCTTCTCTGTTTTCACGTAGTTGGGTTGGGTTGACAAAGTTTCTGCCTTTAAGATAAAT	1943
Db	721	CAAATCATGCTTCTCTGTTTTCACGTAGTTGGGTTGACAAAGTTTTCTGCCTTTAAGATAAAT	780
Qy	1944	GAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTCGTTTCAAG	2003
Db	781	GAGTAATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTTCAAG	840
Qy	2004	AAACAGTAAAACTTGGGTTTCAATCTTTA	2030
Db	841	AAACAGTAAAACTTGGTTTCAATCTTTA	867

LOCUS   BX933564   2064 bp   mRNA   linear   VRT 30-MAR-2004
DEFINITION   Gallus gallus finished cDNA, clone ChEST746b24.
ACCESSION   BX933564
VERSION   BX933564.2 GI:46017940
KEYWORDS   .
SOURCE   Gallus gallus (chicken)
  ORGANISM   Gallus gallus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.

ORIGIN

Query Match   41.0%; Score 859; DB 8; Length 2521;
Best Local Similarity   99.4%; Pred. No. 6.6e-139;
Matches 862; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

REFERENCE   1 (bases 1 to 2064)
AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
TITLE   Direct Submission
JOURNAL   Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT   On Apr 1, 2004 this sequence version replaced gi:41634092.
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from whole embryo, normalised, and poly
            A-trimmed. EcoRI-NotI cut cDNA was then ligated into the vector.
            Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host:
            Escherichia coli DH10B.
FEATURES   Location/Qualifiers
     source   1..2064
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST46b24"
                     /clone_lib="ChEST46b24"
                     /dev_stage="stage 10"

ORIGIN
Query Match   35.6%; Score 745.2; DB 5; Length 2064;
Best Local Similarity   69.8%; Pred. No. 3.7e-119;
Matches 1088; Conservative 0; Mismatches 443; Indels 27; Gaps 5;

Qy   206 CGGCGCAGCCATGTCCCGGGGCGGGCCCGAGGAGGTGAACCGGCTCACGGAGCACCTACC 265
              |||| | ||||| ||||| |||||||| |||||| || | ||| ||||| || |
Db   30 CCGCGCGACCATGTCGCGGGGACCCGGGGAGGAGGTTGACAGCTTAACGGAGAGCACTTACA   89

Qy   266 GGAATGTTATGGAACAGTTCAATCCTGGGCGGCAAATTTAATAAAACTGGGGAGAAAATT 325
              |||||| | |||||||| || ||| ||||| | ||||||||||||| ||||||| ||
Db   90 AAAATGTTATGGAGCAATTCAATCCAGGACTGCCGGAATTTAATAAATCTTGGGAAGAATT 149

Qy   326 ATGAGAAAGCTGTAAAGCGTTATGATCTCAGCAGGAAAAAGCCTACTACGGATGGAGTGGCCA 385
              ||| ||| ||||| | |||||| | |||||| ||||| ||| | ||||| ||||
Db   150 ACGAAAAAGCTGTTAATGCTATGGTCATCGTCGGAAGAATCATACTACGACAGTCTGGGCAA 209

Qy   386 AGATGCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGACCAATGTCCTCATTAG 445
              || || ||||| || ||| |||| || ||||| || ||||||| || |||||| || |
Db   210 AGATTGGGGAATATATCAGCAGAATTCACCTGTTTCTAAAGAGTTAGGCGCAAGTCCTGGTCG 269

Qy   446 AGATTTCAAGTACCCCACAAGAAAAACTTCAGGAGAGTTTCATGTGAAAATTTTAAAAAAATTCC 505
              ||||||||| | ||| || || || |||| || ||||| ||| | ||| |||| ||
Db   270 AAATTTCAAGAACACACAAGAAAAAAAGTTAAGGAGAGAAAAGTTTTTAAGGAAATTTC 329

Qy   506 ACAAAGAGATTATCCATGGAGCTGGAGGAAGAAGATAGAACTTGACGTGAAATATATGAACG 565
              || | |||| | | || ||||||| || || |||||| |||| ||||| |||| | ||
Db   330 ATAAGGAAATTATATCTGAACTGAAGAAAAGAACAGAGCAGAAAAGTAACGACG 389

Qy   566 CAACTCTAAAAAGATAACCAAAACAGAACAGAACAAGAAGAATAAATTAGAGTCTTTGGAGAAATCCC 625
              || | ||| | ||| | || | |||| ||| | | |||| | ||| || ||||||||
Db   390 CCGACATGAATGGCATCGTAAAAACCCCTTTCTCAGGTGGAGAAAATCCTTTGCAACTG 449

Qy   626 AAGCTGGAGTTGAAGAGAAGACTACCCAAAACCAGAAGGAAAGCCGAAAACGACTCAAATATG 685
              |||| ||||| | || ||| | ||| | |||| || | ||| | ||| | |||
Db   450 CGACTTTAAAGAGACTACCAAAACCAAGAAGAGGAAAGTGAAATTTCCCTGGAGAAGTCTC 509

Qy   686 AACACAAAGAAATTGAGTTATGTGGAGGACCGTTACTTCGTCGTGAAAATCCAGAAAAT 745
              || | || || |||| ||| | |||| ||||||||| | || ||||||| |||| |
Db   510 AGCATAAGGAAATGAGTATTTGGAAAACCGTGACGTTCTCGGACAGACTGACATTCAGAGAT 569

Qy   746 TCATTGCAGATGGGTTGCAAAGAGAGGCTCTGCTTGAAGGAAGGAGAGAGAGCGGCCTTCTGCTTTCTGG 805
              || | || | |||| |||||| | ||| | || |||| || ||| | |||| ||||||||
Db   570 TTATTGCTGAAGGCTAGAGGAGAAGTCTACTTTGAAGAAGAGATTCGTGTTTTCTAG 629

Qy   806 TTGATAAGCACTGTGGGCTTTGCAAACCACATAACCACATAAGTTGAAACGGAC 865
              |||| ||||| | ||||||| ||||| | |||| | ||| ||| | || |
Db   630 TTGACAAGCACTGCAGTTTTTACCCAGCACATGCACTTCTATCACGCGAGTGTGGCAGAAT 689

Qy   866 TACTGAAATTCCAAGCTTCGCTCCTCGTGGCGGCGAGGAACGGTTTCACCCAGCCAGTTGCAGTTCCTG 925
              || |||||| ||||| || ||| | ||||| ||| |||| || | ||||||| ||| || | |
Db   690 TCCTCAACGGCCAAGCTGCCCTGGGTGGCAGGAATTTGCAGTGAGTCAGTTCCACCAAAGTTCCTG 749

Qy   926 AGAAAATCATGATATGATCGAAGAAGAAATAAAGACCCAGAGCTTACCCCGTGCTGGAA 985
              ||| | |||| || || |||||||| || |||| ||| || ||| |||||||||||
Db   750 AAAAAGTCAAAATGATTGAGGAGAAAATAAAGGACACCCTGGTTCCAACTCCCATATCCAGGAA 809

Qy   986 CTCCTCAGGCGTTCACCCATGATCGGATCGGAGAGAGGAGCACAATGTGGTTAGGAGAAAGATTACGGACACCC 1045
              || | ||||| || |||| || || ||||| || ||||||| | || ||||| ||| | | || |||||||
Db   810 ATAATGAAAATGCATCAAAAAGGTGCCCCCTCTGTGGGCGGGGGCATAACCACCAGGGC 869

Qy   1046 TTTCTAAATGCTCACCAAAAGAATGCCAGAAAGAGGTGCCCCCGTCCTCCTTCAGGCAGCAGACCATATAACCAGTCCCT 1105
              ||| ||| |||| || |||||| | ||| | | ||| || | ||||| |||||| |||| ||| || || |
Db   870 ATAATGAAAATGCATCAAAAAGGTGCCCCCTCTGTGGGCGGGGGCATAACCACCAGGGC 929

Qy   1106 TGATCGGATATGTTTAATAACCCAGCCACGGCTGCCCCCGGAATTCAC---AAAGGGTAAATA 1162
              ||| | | | || ||||| | |||| | | ||| | | |||||| ||||| ||
Db   930 TTGTGGACATGTTCAACAATCCTCCCATGGCTCTCCAAAAGGTGTCTCCCTGAAAGACCGACCA 989

Qy   1163 ATTCAACAGGTACTTCCGGAAGATCCCAGTGTTACAGCGCAGTCAGTTTTGACTGCTGTTTCGGTTGCAACGGGGAC 1222
              ||| | |||| ||| | ||| |||||||||| || | ||| ||||| | | | || |||| |||| | |||
Db   990 ACCCAGCAGAGAATTCAGATGATGACCAGTTTATCACGCCTCAACTTCCTGTTCCAACAGGGC 1049

Qy   1223 TGAACATGATGAACAGAAGACAGAAAGTGAAGAACCATCTTCCGCGACACTGCGGGGCCTCCAACA 1282
              ||| | |||||| ||| | | | | |||| || | ||| ||||||||| || | |||| ||| |
Db   1050 TAAATATAATGAAAAAGGCAGAGAAGTAAGGACAATCTTTCCACATATCTGCTGCATCATCAAGCGCTG 1109

Qy   1283 AGACCTTACTCAGCTTTGCACTGCCACGGCTGCTCATCCCCGAGGAGAAGG 1342
              |||| ||| ||| |||| |||| ||||| ||||| |||||||||| |||
Db   1110 AGACATTACTTAGCTTTTGCACAGCCACCATCACTCTTGTATCTGAAGAAAAG 1169

Qy   1343 ATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGCGGAGGGGTTGGTTGGTTCCCGTCGTT 1402
              |||| | ||| | ||||| || | | | |||| || || ||| ||| ||||||| | | |
Db   1170 ATGGCTGGCTTTACGGGACATCAACTAAATGGAGTTTTTCCATCATCAT 1229

Qy   1403 ACACGAAAGTTGCTGAAAGAAAAATGAGACAGAGTGACCGTGCCCAGCGCCAAGCCCCA 1462
              || |||| ||| | ||| | | | ||||| || || || |||| |||| ||| ||
Db   1230 ATACCAGACACAGTAGAAGAATCAGA---AGAAAAAGCCTTTGTTCTGGAGCCATCAAGGCGG 1286

Qy   1463 CACCAGTGAGAAGCATCAGACCCGTGAACTTGTCTGAGAATAGCACGTGTGTCATCCCC 1522
              ||||||| |||| ||||| | ||| || | | |||||| | ||||| ||| ||||
Db   1287 CGCCGGTTAGGAGTATAGCTCAGAATCTGAAATTCTGCAGAAAAAGCTGGTTGCTCCCCAC 1346

Qy   1523 CACCCGGACTACTTGGAAATGCTTGTCCATGGGGGGCAGGGCTGCGACAGGAGCAGAATTCGG 1582
              || || ||| | | ||||| || | | |||||| |||| |||| | | | || ||
Db   1347 CGCCAGACTATCTGGTT--------CAACTTCAGATAAGAGATGGAGTCTT 1391

Qy   1583 CCAGGACGACATCCACCCTTTAAGGCCCGGAGACCGCCTCCTAA--1640
              ||||| | | |||| | ||||||| | || || ||| ||
Db   1392 CCAAAGGTACTTGCTGC---TAGAACACAACCTGACAAAACAGAAAGTTCTGAAAC 1448

Qy   1641 -CGATGCCAACGGGACTGCAAAGCCGCCTTTTCTCAGGGGAGAAAAGCCCTTTGCCACTG 1699
               |||| | |||| | ||||| |||| |||| | |||| | | |||||| | |||||
Db   1449 CCGGACATGAATGGCATCGTAAAAACCACCCTTTTCTAAGTGGAGAAAATCCTTTGCAACTG 1508

Qy   1700 TGAAACTCCCGCCCGACTGTGACGTAGGATCGGTCGGATGGATGAGGAGGA 1757
              ||||| || | | | |||| | || ||| || || | || |||||||
Db   1509 TGAAACTCAGACCAACAGTAACAAATGACAGATCCACCATTATTCGGATGAACACAGA 1566

RESULT 10
AK074419
LOCUS   AK074419   2296 bp   mRNA   linear   PRI 12-SEP-2003
DEFINITION Homo sapiens cDNA FLJ23839 fis, clone KAT02564.
ACCESSION   AK074419
VERSION   AK074419.1 GI:18677016
KEYWORDS   oligo capping; fis (full insert sequence).

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2296)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Suzuki,Y., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:ficdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5416,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction;
            5'- & 3'-end one pass sequencing: Department of Virology and Human
            Genome Center, Institute of Medical Science, University of Tokyo
            (partly supported by Science and Technology Agency).
FEATURES             Location/Qualifiers
     source          1..2296
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAT02564"
                     /cell_line="KATO III"
                     /cell_type="signet-ring cell carcinoma"
                     /clone_lib="KAT"
                     /note="cloning vector: pME18SFL3"
ORIGIN
Query Match       32.0%;  Score 671;  DB 8;  Length 2296;
Best Local Similarity   100.0%;  Pred. No. 2.8e-106;
Matches   671;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1360 GAACACGAGCGTGTCCAAGGGCGAGGAGGGGGTTGGTTCGGTGTCGTTACACGAAGTTGCTGGAA 1419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACACGAGCGTGTCCAAGGGCGAGGAGGGGGTTGGTTCGGTGTCGTTACACGAAGTTGCTGGAA 60

Qy       1420 GAAAAATGAGACAGAAGCACAGTGACCGTGCCCACACGCCCAAGCCCCCCACACAGTGAGGCATC 1479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAAATGAGACAGAAGCACAGTGACCGTGCCCACACGCCCAAGCCCCCCACACAGTGAGGCATC 120

Qy       1480 AGCACCCTGAACTGTGTCGAGAATAGGACAGTGTGTCATCCCCGACCTACTTGGAA 1539
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACCGTGAACTTGTCTGAGAATAGGACAGTGTGTCATCCCCGACCTACTTGGAA 180

Qy       1540 TGCCTTGTCCATGGGGGAGCTCCGACGAGGACAGATTCGGCCACGACGACATCCACC 1599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCCTTGTCCATGGGGGAGCTCCGACGAGGACAGATTCGGCCACGACGACATCCACC 240

Qy       1600 TTTTAAGGCCCCCAGCGGTCCAAGCCCGAGCCGCGGCGCTCCTACGATGCCAACGGGACTGCA 1659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTAAGGCCCCCAGCGGTCCAAGCCCGAGCCGCGGCGCTCCTACGATGCCAACGGGACTGCA 300

Qy       1660 AAGCCCGCCCTTTTCTCAGCGGAGAAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTG 1719
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGCCCGCCCTTTTCTCAGCGGAGAAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTG 360

Qy       1720 ACGAATGATACGGCTCGGCGCCACCCATCATTCGATCAGAGAGACCACGCCAAGGACTCTCCCGCGGCC 1779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACGAATGATACGGCTCGGCGCCACCCATCATTCGATCAGAGAGACCACGCCAAGGACTCTCCCGCGGCC 420

Qy       1780 TCTCCGGTTCCTCCCTTGCGGAATGATGGGGGGCGCATCCTGTCTGCTCGCCACGTGCTGACGGTCGG 1839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTCCGGTTCCTCCCTTGCGGAATGATGGGGGGCGCATCCTGTCTGCTCGCCACGTGCTGACGGTCGG 480
```

```
LOCUS       BC061676                3616 bp    mRNA    linear   VRT 03-FEB-2004
DEFINITION  Xenopus laevis hypothetical protein MGC68822, mRNA (cDNA clone
            MGC:68822 IMAGE:4633564), complete cds.
ACCESSION   BC061676
VERSION     BC061676.1  GI:38197625
KEYWORDS    MGC.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 3616)
  AUTHORS   Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
            and Richardson,P.
  TITLE     Genetic and genomic tools for Xenopus research: The NIH Xenopus
            initiative
  JOURNAL   Dev. Dyn. 225 (4), 384-391 (2002)
  PUBMED    12454917
REFERENCE   2  (bases 1 to 3616)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   3  (bases 1 to 3616)
  AUTHORS   Klein,S. and Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2003) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
REMARK      NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
            cDNA Library Preparation: Life Technologies, Inc.
```

```
Qy       1840 GAAGCTTCAGTGGAGGAGGCCTAACTCTAATGTCGCCGCCGCTTAAGCAAATACATGCTTCTCT 1899
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGCTTCAGTGGAGGAGGCCTAACTCTAATGTCGCCGCCGCTTAAGCAAATCATGCTTCTCT 540

Qy       1900 GTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATG 1959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATG 600

Qy       1960 ACCAGCTCAGCCATTTAAAATATTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGT 2019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGCTCAGCCATTTAAAATATTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGT 660

Qy       2020 TTCAATCTTTA 2030
              |||||||||||
Db        661 TTCAATCTTTA 671
```

```
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: National Institutes of Health Intramural
                  Sequencing Center (NISC);
                  Gaithersburg, Maryland;
                  Web site: http://www.nisc.nih.gov/
                  Contact: nisc_mgc@nhgri.nih.gov
                  Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
                  Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
                  Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
                  Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
                  Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
                  McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
                  Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
                  Young,A., Zhang,L.-H. and Green,E.D.

                  Clone distribution: MGC clone distribution information can be found
                  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
         Series: IRAK Plate: 128 Row: P Column: 13.
FEATURES          Location/Qualifiers
     source       1..3616
                  /organism="Xenopus laevis"
                  /mol_type="mRNA"
                  /db_xref="taxon:8355"
                  /clone="MGC:68822 IMAGE:4633564"
                  /tissue_type="lung, adult Xenopus"
                  /clone_lib="NICHD_XGC_Lu1"
                  /lab_host="DH10B"
                  /note="Vector: pCMV-SPORT6"
     gene         1..3616
                  /gene="MGC68822"
                  /db_xref="GeneID:398987"
     CDS          115..1641
                  /gene="MGC68822"
                  /codon_start=1
                  /product="MGC68822 protein"
                  /protein_id="AAH61676.1"
                  /db_xref="GI:38197626"
                  /db_xref="GeneID:398987"
                  /translation="MSRDADAVAKITESTYRNVMNQFNPGLRNLVNLGRNYEKAVAAM
                  VQAGRAYYDGVAKMGELASASHVSKELGHVLLBISIVHKRLMDNLEENFKFHREIVA
                  ELERKTDQDIKYMNATLKRYQTEHKNKVDSLEKSQSDLKKLRKSQGGRNVTKYEIKE
                  SEVETITSRQSEINKFIAQGCREALLEEKRFCFLVDKHCTFSNYINYFHVQAVDLL
                  TSKLPKWQEOTCSDISQVPDTVVNIMQDLKTPVSTPISASPEFSPLMDKKALMAGDTLY
                  RNTLRMPPAPQTKAQTSPLVDMFNKPEVVQKLTFDNPRTSQESQSEGTLPRSMSVAT
                  GLNOMKRKVKTIFPPHSGNGSTQMSFSQGDIITLLIPEEKDGWLYGEHDTTKVKGWFP
                  SSYTRPLEENGREPLKVPSVSPSPAPVRSLSTANLVEKSGVVLPEPDYLEPMSGKRSS
                  SNNSNRGYTFSPLIAKDPPSTANGITNHPLSGENFFSTILRPTVTNDRSAPIIR
                  "
     misc_feature 1132..1311
                  /gene="MGC68822"
                  /note="SH3; Region: Src homology 3 domains"
                  /db_xref="CDD:smart00326"
ORIGIN
```

```
Best Local Similarity   28.4%;
Matches 1014; Conservative   0; Mismatches 511;   Indels   33;  Gaps    8;
```

```
Qy    208  GCCGCAGCCATGTCCCCGGGGGGCCCCAGGAGGTGAACCGGCCTCACCGAGACACCTACCGG  267
                ||      |||||| |     |||||              |  |   ||  |||||  |
Db    106  GACGTAGCCATGTCCCGGACGCAGAGCCGTGGCAAAGATCACAGAGAGCAGCGTACAGG  165

Qy    268  AATGTTATCGGAACAGTTCAATCCTGGGCTTAATAAACCTGGGCGAAAATTAT  327
                | |||| |||  || |  ||||  ||| ||||    ||||           |||
Db    166  AATGTTATGAATCAGTTCAACCCTGGGTTAAATCTTGGAAGGAACTAT  225

Qy    328  GAGAAGAGCTGTAAACGCCTATGATCCTGGCAGGAAAAGCCTACTACGGATGAGTGGCCAAG  387
                ||  ||| ||         || |||      ||||| |||||||  ||   | ||||||
Db    226  GAAAAAGCAGTGGGCAGCAATGGTGCAGGAAGCGGTATTATTATGATGACCAAG  285

Qy    388  ATCCGTGAGATTGCCACTGGGTCCCCCCGTGTCAACTGAACTGGGACACATGTTCCTCATAGAG  447
                |  |    ||    |||  |||| |    ||||  ||    ||  || | |||  || |
Db    286  ATGGGCGAATTGGCCTCTGCATCTCATGTATCAAAAAGAGCTAGGCCACGTACTTCTGGAG  345

Qy    588  ATTAAAGGAATCCGGAGTATGTGGGAAAACCAATAAGCTCAGACACAGTGCGGAGTAAAATTC  645
Db    586  ...
```

```
Qy    748  ATTGGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAGAGAGAGAGCTGCCTTTCTGGTT  807
Db    646  ATAGGCGCAGGGGGCTGCAGGGGAGGCTTCGGCTTGAGGAGAGAGAGACGAATTCTGTTTCTTTAGTG  705

Qy    808  GATAAGCACTGTGGGCTTTGCAAACCACATACATTATTATTCACATTACAGTCTGCAGAACTA  867
Db    706  GACAAACACTGCACCTTCTCGAATTATTATCAACTATATTTCCACGTACAGCGCTGTCTA  765

Qy    868  CTGAATTCCAAGTCTGCCTCGGTGTGGCGAGAAGGACCTGTGTTGATGCGCCATCAAAAGTCCAGAG  927
Db    766  CTCACCTCCAAGTCCCCAAGTGGCAGGGCCGCAGAGAGCCTGTGTGAGCATCTCTCAGGTTCCAGAC  825

Qy    928  AAAATCATGAATATGATCGAAGAAATAAGAGACCCCAGGCAGCCTTCTCACCCCGTGTCTGGAACT  987
Db    826  ACAGTCGTTAATATTATGCAGGACCTGAAAAACTCCAGGGTGTCCACACCAATCTCTGCTTCC  885

Qy    988  CCTCAGGCTTCACCCATGATCGAGAGAAGCAAATGTGGTTAGGAAAGATTACGACACCCTT  1047
Db    886  CCTGAACCATCCCCATTAATGGA-----------CAAAAGGCCCTGATGGCTGGAGATACCCTG  939

Qy   1048  TCTAAAATGCTTCACCAAAAGATGCCCCCCGTCTCCCTGAGGCAGAGAGCGAATATACCAGTCCCTTG  1107
Db    940  TATAGGAACACACTAAGGATGCCTCCAGCACCCCAGACGAAAGCACGAGTACCCTG  999

Qy   1108  ATCGATATGTTTAATAACCCAGCAGCCACGGCTG-----CCCGAATTCACAAAGGGTAAATA  1162
Db   1000  GTTGACATGTTTAACAAAACCACCAGTTGTCCAAAAACTAACATTTGACAATCCAAGGACA  1059

Qy   1163  ATTCAA-CAGGTACTCTCCGAAGATCCCAGTTTACAGCGGATCCAGTTTCGTTGCAACGGGA  1221
Db   1060  TCACAAGGAGAGAGTCAAGAAGAGGCACTTTACCTGCGTCCATGTACGTAGCAACAGGT  1119

Qy   1222  CTGAACATGATGAAGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGTCCAAC  1281
Db   1120  TTGAATCAAATGAAAAAAGCAGAAAAGTCAAAACCATCTTCCCACACA----GCGGAAAGGC  1176

Qy   1282  AAGACCTTACTCAGCTTTGCACGGGAGATGTCATCACGGTGCTCATCCCGGAGGAG  1341
Db   1177  AGCACACAGATGAGCTTTTCTCAGGGGCAGCACTTCTTATCCCAGAGGAAAG  1236

Qy   1342  GATGGCTGGCTCTATGAGAACACACGTGTCCAAGGGGACGGGGGTTGGGTTCCCGTCG  1401
Db   1237  GATGGCTGGCTGTACGGGGAGCACATGATAACCAAAGTTAAAGGTTGGGTTCCCATCATCC  1296

Qy   1402  TACACCGAAGTTGCTGGAAGAGAAAAAATGAGACAAGAGCAGTGACCGTGCCCA-----CGCCA  1455
Db   1297  TATACAAGGCCGCTGGAAGAAAATGGCCGAAGGTGCCAAGGCGTGTCCCA  1356

Qy   1456  AGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTC  1515
Db   1357  AGTCCTGCTTCCAGTCCGGAGTCTCAGCCAGCGGCCAACTTGTGAAAAAGCCAGAGTGTC  1416
```

```
RESULT 12
LOCUS       AX379210               559 bp    DNA     linear   PAT 18-MAR-2002
DEFINITION  Sequence 252 from Patent WO0196389.
ACCESSION   AX379210
VERSION     AX379210.1  GI:19575050
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Meagher,M.J., King,G.E., Xu,J. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196389-A 252 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match             26.5%;  Score 556;  DB 6;  Length 559;
Best Local Similarity   99.5%;
Matches  556;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;
```

```
Qy      1516 ATCCCCCCACCCGACTACTTGGAATGCTTGTGTCATGGGGGGCAGCAGCTGCCGACAGGAGAGCA 1575
             |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1417 CTCCCAGAACCGGATTATTTGGGAACCAANTGTC--TGGGAAAACCAGTAGCAGCAGCAACA 1474

Qy      1576 GAATTCGGCCCAGGACGACATCCACCTTTAAGGCCCCCCAGCCTTCCAAGCCCGAGACCCGCGGCT 1635
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1475 ATAGCAACCGGGGGGATACA-----CTTTCAGCCCAC-----TCAATTGCCAAAGATCCTCCT 1524

Qy      1636 CCTAACGATGCCAACGGGACTGCAAAGCCGCCTTTTTCTCAGGCGGAGAGAAAACCCCTTTGCC 1695
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1525 CCTTCAACTGCAAAATGGCATCACTAACCATCCATTTCTAAGTGGAGAAAAACCCATTTTCC 1584

Qy      1696 ACTGTGAAAACTCCGCCCCGACCTGTCAGGAATGCATGATCGGCGGCCACCCCATCATTCGATGAG 1753
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1585 ACCATCAAAGCTCCGACCAACCGTAACAAAATGACGACCTGCGCGCCAATCATCCGATGAG 1642
```

```
RESULT 13
LOCUS       BX931244               685 bp    mRNA    linear   VRT 02-FEB-2004
DEFINITION  Gallus gallus finished cDNA, clone ChEST261d10.
ACCESSION   BX931244
VERSION     BX931244.1  GI:41631772
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.B.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Graham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bbs.umist.ac.uk
COMMENT     BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from ovary, normalised, and poly A-trimmed.
            EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
            pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
            coli DH10B.
FEATURES             Location/Qualifiers
     source          1..685
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer"
                     /db_xref="taxon:9031"
                     /clone="ChEST261d10"
                     /clone_lib="CSEQRBN21"
                     /dev_stage="adult"
ORIGIN

Query Match             18.9%;  Score 397;  DB 5;  Length 685;
Best Local Similarity   73.7%;       Pred. No. 9.2e-59;
Matches  505;  Conservative   0;  Mismatches 180;  Indels   0;  Gaps   0;
```

```
Qy      1283 AGACCTTACTTCAGCTTTGCACAGGGAGGAGAGTGTCATCAGCGCTCATCCGCTCAT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AGACCTTACTCAGCTTTGCACAGGGAGGAGATGTCATCAGCGGCTTCATCCCGAGGAAGG 480

Qy      1343 ATGGCTCGGCTCTATGGAGGAGAACACGACGGTGTCCAAGGGAGGGTTTGGTTCCCGTTCGTCGT 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ATGGCTGGGCTCTATGGAGGAGAACACGACGGTGTCCAAGGGAGGGTTTGGTTCCCGTTCGTCGT 540

Qy      1403 ACACGGAAGTTGCTGGAAGA 1421
             ||||||||||||||||||||
Db       541 ACACGGAAGTTGCTGGAAGA 559
```

```
Qy       863 AACTACTGAATTCAAGCTGCTTCGCTTCCGGTGTTGGGCAGGGGGCATCAAAGTGC 922
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AACTACTGAATTCCAAGCTGCCTCCGCNGGCAGGAGACCTGTTGTTGATGCCATCAAAGTGC 60

Qy       923 CAGAGAAAATCATGAAATATGAGATTGATCGAAGAAAATAAAGAGCCCTCTACCCGCGTGTCTG 982
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CAGAGAAAATCATGAAATATGATCGAAGAAATAAAGACCCTCTACCCGCGTGTCTG 120

Qy       983 GAACTCCTCAGCGCTTCACCCATGAGAGAGCAAATGTGCGTTAGAAAAGATTACCACA 1042
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GAACTCCTCANNCTTCACCATGATCGAGCGAAGCAAATGTGGTTAGGAAAAGATTACGACA 180

Qy      1043 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCCGCCTCCTTCAGGCAGCAGCAATATACCAGTC 1102
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCCTTTCTAAATGCTCACCAAAGATGCCCCCCCGCCTCCTTCAGGCAGCAGCAATATACCAGTC 240

Qy      1103 CCTTGATCGATATGTTAATTAATAACCAGCAGAATTCACAAAGGGTAAATA 1162
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CCTTGATCGATATGTTAATTAATAACCAGCAGAATTCACAAAGGGTAAATA 300

Qy      1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGGGATCAGTTTGCGACGGGAC 1222
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGGGATCAGTTTGCAACGGGAC 360

Qy      1223 TGAACATGATGAAGAAGAGCAGAAAAGTGAAGACCATCTTCCCGCACACTGCGGGGCTCCAACA 1282
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGAACATGATGAAGAAGAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGGCTCCAACA 420
```

```
Qy       174 CCAGCCGCCGTCTCTCCCCCGCGCGTTGCTCTCCCCGCGGCCGCCAGCCGCAGGGGCGGGGCCCGA 233
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CCCGGAAGGCGCTCTGCGGAGAGAACGAGGCGTGTCCCCGCGCCGCCCAGCCGCAGGGGGGGCCCGA 60

Qy       234 GGAGGTGAACCGGCTCACGGAGGAGCACCTACCGGAANGTTATGGAGAACAGTTCAATCCTCTGG 293
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GGAGGTGAACCAAGCTACCGGAGGAGCACTTACAAAAAATTATGAGGAAAGCTGTAAACGCTATGATCGG 120

Qy       294 GCTGGCGAAATTTAATTAAAACCTGGGGGAAAAATTATGAGGAAAGCTGTAAACGCTATGATCCT 353
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ACTGCCGAAATTTAATAAAATCTTGGGAAGAAATTACGAAAAAGCTGTTAATGCTATGGTAGT 180

Qy       354 GGCAGGAAAAGCCTACTACGATGGAGGTGGCCAAGATCGGTGAGATTGGCCACTGGGTCCC 413
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 14

LOCUS   BC068330   2563 bp   mRNA   linear   VRT 30-MAR-2005
DEFINITION   Danio rerio BAI1-associated protein 2-like 1, mRNA (cDNA clone
             MGC:85624 IMAGE:6525249), complete cds.
ACCESSION   BC068330
VERSION   BC068330.1 GI:4649678
KEYWORDS   MGC.
SOURCE   Danio rerio (zebrafish)
  ORGANISM   Danio rerio
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
             Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 2563)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2 (bases 1 to 2563)
  AUTHORS   NIH MGC Project
  CONSRTM   NIH MGC Project
  TITLE   Direct Submission
  JOURNAL   Submitted (02-APR-2004) National Institutes of Health, Mammalian
             Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Sequencing Group at the Stanford Human Genome
             Center, Stanford University School of Medicine, Stanford, CA 94305
             Web site: http://www-shgc.stanford.edu
             Contact: (Dickson, Mark) mcd@paxil.stanford.edu
             Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
             R. M.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAK Plate: 165 Row: b Column: 19
             This clone was selected for full length sequencing because it
             passed the following selection criteria: matched mRNA gi: 38708008.

FEATURES   Location/Qualifiers
     source   1..2563
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="MGC:85624 IMAGE:6525249"
                     /tissue_type="Kidney, zebrafish"
                     /clone_lib="NCI_CGAP_Zkid1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6.1"
     gene   1..2563
                     /gene="baiap2l1"
                     /gene="synonyms: cb1023, zgc:85624, loc55971"
                     /db_xref="GeneID:387261"
                     /db_xref="ZFIN:ZDB-GENE-031201-1"
     CDS   139..1596
                     /gene="baiap2l1"
                     /codon_start=1
                     /product="baiap2l1 protein"
                     /protein_id="AAH68330.1"
                     /db_xref="GI:4649679"
                     /db_xref="GeneID:387261"
                     /db_xref="ZFIN:ZDB-GENE-031201-1"
                     /translation="MSRASEDVNKLTSTYKNVMEQFNPGLRNLVNLGKNYEKSVAAM
                     TFAGKAYFDALSKIGESAAVSPVSRELGVVLMEISEVHKNAQLEMESFKRFHKELIA
                     ELEKKTDMDTKYKMTATFRYQSEHKLKQDFLDKSQADLKKRKSQGKHSSKYEIKEN
                     ECMETISSRQTDMQRFIADGCKEALLEEKRRPCFLVDKICALSYQLSAFHDKAKEMLT
                     LKLPSWQEKCNDANRVPDTVISMIEGLRTSMSVIPESSPTLERSDRTNSGSIVPPPAP
                     SLKAHTSPLASMPSQEBAPKSPISSEQSYSDQDSLDENGLSRSTVSSGMDTAKKTRVQT
                     IFPHTAGNNETLLSFEDGDIIIILIQEERDGMIYGELEHTGQRGWFPSSYCRSYNETT
                     IPNSRMESPFSSHIDLPEQOEBDEEVEBPVLLPPPDYSDSSRDRPIKPSTPSSQSTVS
                     LANGHSMPPFLGGGNPFATVRLRPTVINDRSAPII"

ORIGIN

Query Match   18.3%; Score 382.6; DB 5; Length 2563;
Best Local Similarity   60.0%;
Matches 712; Conservative   0; Mismatches 439; Indels 36; Gaps 3;

Qy   217 ATGTCCCGGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGGAGCACCTACCGGAAGTTATG 276
           |||| |||| | |||||| || |||| || ||| ||| || || || || || |||
Db   139 ATGTCTCGAGCCTCGGAGGACGTGAACAAACTCACCGGAGAGCACATACAAGAGAATGTGATG 198

Qy   277 GAACAGTTCAATCCTGGGCTGCGGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCT 336
           |||||||| ||| || ||||| ||||| ||||| |||| || ||||| |||| ||||| |
Db   199 GAACAGTTCAACCCAGGACTGCGGAATCTCGTTAACCTGGTTAACCTATGAGAAAATCA 258

Qy   337 GTAAAGCTATGATCCTGGCAGAAAAGCCTACTACGGATGGGAGTGGCCAAGATCCGTGAG 396
           || || |||||| ||||| |||||||| ||||| || ||||| |||| ||| ||| |||
Db   259 GTTGCAGCAATGACCTTTGCCGGAAAGGCTTATTTGATGCACTCTCAAAAATAGGGGAA 318

```
RESULT 15
LOCUS       AY398378                2566 bp    mRNA    linear   VRT 16-NOV-2004
DEFINITION  Danio rerio clone RK125A2C09 insulin receptor tyrosine kinase
            substrate (LOC55971) mRNA, complete cds.
ACCESSION   AY398378
VERSION     AY398378.1  GI:37681866
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 2566)
  AUTHORS   Song,H.D., Sun,X.J., Deng,M., Zhang,G.W., Zhou,Y., Wu,X.Y.,
            Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zon,L.I.,
            Kanki,J.P., Liu,T.X., Look,A.T. and Chen,Z.
  TITLE     Hematopoietic gene expression profile in zebrafish kidney marrow
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (46), 16240-16245 (2004)
   PUBMED   15520368
REFERENCE   2  (bases 1 to 2566)
  AUTHORS   Zhang,G.W., Sun,X.J., Wu,X.Y., Song,H.D., Zhou,Y., Liu,T.X.,
            Deng,M., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y.,
            Zon,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2003) State Key Lab for Medical Genomics,
            Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
            Shanghai Second Medical University, 197 Rui Jin Road II, Shanghai
            200025, P. R. China
FEATURES             Location/Qualifiers
     source          1..2566
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="RK125A2C09"
                     /tissue_type="kidney marrow"
     gene            1..2566
                     /gene="LOC55971"
     CDS             156..1613
                     /gene="LOC55971"
                     /codon_start=1
                     /product="insulin receptor tyrosine kinase substrate"
                     /protein_id="AAQ97811.1"
                     /db_xref="GI:37681867"
                     /translation="MSRASEDVNKLTESTYKNVMEQFNPGLRNLVNLGKNYEKSVAAM
                     TPAGKAYFDALSKIGESAAVSPVSRELGVVLMEISEVHKNAQLEMEESFKRFHKELIA
                     ELEKKIDMDTFKYMTATFKVQSEHKLHKQDFLDKSQADLKKLKRRKSQGKHSSKYEIKEN
                     ECMETISSRQTDMQRPIADGCKEALLEEKRRPCFLVDKICALSYQLSAFHDRAKEMLT
                     VKLPSWQEKCNDATRVPDTVISMIEGLRTSMSVIPESSPTLERSDRTNSGSIVPPAP
                     SLXAHTSPLASMFSQEAPKSPISSEQYSDQSLDENGLSRSTSVSSGMDTAKKTRVQT
                     IFPHTAGNNETLLSFEDGDIIILLIQEERDGWLYGELEHTGRGWFPSSYCRSYNETT
                     IPNSRMESPFSSHIDLPEQOEEDEVEEPVLLPPPYSDSSRDRPIKPSTPSSSQSTVS
                     LANGHSMPFLGGGNPFATVRLRPVTVNDRSAPII"
ORIGIN
Query Match        18.2%;   Score 381;  DB 5;   Length 2566;
Best Local Similarity   59.9%;   Pred. No. 6e-56;
Matches  711;  Conservative   0;  Mismatches  440;  Indels   36;  Gaps   3;
```

```
Qy    397 ATTGCCACTGGGGTCCCCCGTGTCAAACTGGAACTGGGACAATGTCCTCATAGAGATTTCAAGT 456
          |||| ||| ||  |  |  || ||||| |  ||| ||  |||| ||  ||||||||||| |||||
Db    319 AGTGCTGCTTGTTTCGCCCGGTTTCTTAGGGAGCTGGGAGTGGGTTAATGGAGATTTCAGA 378

Qy    457 ACCCACAAGAAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATT 516
          ||    ||||||  |  | ||||||||  | |||||||| |||||||||| |||||||| ||
Db    379 GTGCACAGAAACGCCCAATTGGACGAGAGGTCTTCAAGACGTTCCACAAAGAGCTC 438

Qy    517 ATCCATGAGCTGGAGAAGAAATTGGACATGGATACCAAGTACTACATGAAGGCCAACTCAA 576
          |||||| |||||||| ||||| ||||| || ||| | |||| ||| | |||| || |  | |
Db    439 ATTGCTGAACTCGAAAAGAAAATCGATATGGATACCAAGTACTACATGAAGGCCAACTCTAAA 498

Qy    577 AGATACCAGTCAGACACAAATTAAAGCAAAGAAGACTTCCTGGACAAATCACAGGGCTG 636
          |||| | |||| || | || | |||||||  |||| || |||| | || | ||||| ||
Db    499 AGATACCAGTCAGACACAAATTAAAGCAAAGAAGACTTCCTGGACAAATCACAGGGCTG 558

Qy    637 AAGAAGATCAGGAAAGGCCAAGGAAAGGCCACTCAAATATGAACACAAGA 696
          ||||||| | | | | | |  || ||| | | |||||||| ||||  ||||
Db    559 AAGAAAACTACGAAGAAAAGTCAAGGA---AACATTGCTCAAATATGAGATCAAAGAA 615

Qy    697 ATTGAGTATGTGGAACGGTTACTTCTCGGTCAGAGAACTACTGAATTCC 756
          |||||| |  | ||  | ||  | |  ||  || ||||  | ||||||||
Db    616 AATGAGGTGCATGGAGACCATCTCATCCCGGCAGACAGCACATGGAAATCCAGAAAATTCATTGCAGAT 675

Qy    757 GGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGAGCCGTTGCCTTCGTGCTTTCTGCTTTCTGGTT 816
          |||| | ||||| | | |  |||| | || || || |  || | |||||||| | || | |||||
Db    676 GGAATGTAAGGAGGCTTTACTTGAAGAGAAAAGGAGAGCGTTTTGTTTCCTGGTGGACAAACAC 735

Qy    817 TGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 876
          ||||| |  | ||  | | | |  | ||| || ||||||  ||||  |   ||||  | |
Db    736 TGTGCCCTTTCTTATCAACTATCAGCTTTCCATGACAAGGGCCAAGGAGATGCTGACCGTC 795

Qy    877 AAGCTGCCTCGGTGGCAGGAGAAGTGTTGATGATGCCATCAAAAGTGCCAGAGAAAATATG 936
          |||||||| ||||| || || ||  | ||||| || || ||||||| ||||| ||| ||
Db    796 AAAACTGCCCAGCTGGCAGAGAAAAGTGCAATGATGCCAACCAGAGCTCCACCACGTGATA 855

Qy    937 AATATGATCGAAGAAATAAAGACCCCAGCCTTCACCCCCGTGTCTTGAAACTTCCTCAGGCT 996
          ||| || ||||| |  || || ||| |  ||| |    || | |  | |||||  ||||| |
Db    856 TCCAATGATTGAGGGGTCTGCGGACTTCA------ATGTCAGTGTACCACAGAGTCT 903

Qy    997 TCACCCATGATCGAGAGAGAAGCAAATGTGGTTAGGAAAGAATTACGGACACACCCTTTCTAAATGC 1056
          |||||    | | || || ||  | ||  | |  || ||   | || || | |  | ||| | || |
Db    904 TCACCTACTACTGGAACGGCAGT--------GACAGGACCAATTCGGTCGGGTC 942

Qy   1057 TCACCAAAGATCCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATG 1116
          ||||| ||| | | | | ||| | |||| |||| |||| || |||| |  |  ||| |
Db    943 TCAATTGTTCACCGCCAGCCCCATCACTTAAAAGCCCACAGCCTGCCCACATG 1002

Qy   1117 TTTAATAAACCCAGCCAGGGTGCCCCGAATTCACAAAAGGGTAAATAATTCAACAGGTACT 1176
          ||   | |||  | |    | |  | |||  || || ||  |  |  ||  ||| ||  | |
Db   1003 TTCTCCCAGGAAGCTCCCAAAAGCCCCATCTCATCAGAGCAGTATTCAGATCAGAGCCAGT 1062

Qy   1177 TCCGAAGATCCCAGTTTACACGGATCCAGTTTGCAACGGCACTGAACATGATGATGAAG 1236
          |||| |  | || ||  |  | || | ||  ||  || |  | || ||| |   || ||
Db   1063 CTAGATGAAAACGGGCTGTCTCGGTCTACGTCAGTGTCTTCGGGTATGGACACGGCAAG 1122

Qy   1237 AAGCAGAAAGTGAACAGCATCATCTTCCCGCACACTGCGGGGCTCCAACAAGACTTACTCAGC 1296
          ||||  ||  |  | | ||| ||||| ||| || | || ||| | ||| | ||| || | ||
Db   1123 AAGACCAGGGTGCAAACTATTTTCCCACACACTGCGGGGAATAACGAAACGCTGCTCAGC 1182

Qy   1297 TTTGCACAGGGAGATGTCATCACGCTGCTCATCCCCGAGGAGAAGAGGATGGGCTGGCTCTAT 1356
          ||||| | || ||| | || ||  || |  |  |   || ||||||| |  |||||||| |||
Db   1183 TTTGAGGATGGAGACATCATTATTTTGCTTATCCAGGAAGAGCGGGACGGGTGGCTGTAC 1242

Qy   1357 GGAGAGCTCGAACATACAGGAAGGGGATGGTTTCCCGTCGTCGTA 1403
          |||| | || || | || || |  || ||||| || |  | | ||
Db   1243 GGAGAGCTTGAACATACGGGACGCGGGTGGTTTCCCTCTTCTTA 1289
```

```
Qy    217 ATGTCCCGGGCGGGGGGCCCGAGGAGGTGAACGGGCTCACGCAGGAGCACCTACCGGATGTTATG 276
          |||||  ||  |  |||| ||| |||| ||||  ||| ||| || || || ||| ||| || |
Db    156 ATGTCTCGAGCCTCGGAAGACGTGAACAAACTCACCGAGAGCACATACAAGAATGTGATG 215

Qy    277 GAACAGTTCAATCCTGGGCGTCTGAGAAATTTAATAAACCTGGGGAAAAATTATGAGAAGTCT 336
          ||||||||||| || || || || ||   || | |||  | || || || | || || |
Db    216 GAACAGTTCAACCCAGGACGCCTGAGAAATCTGGTTAACCTGGGGAAGAACTATGAGAAATCA 275

Qy    337 GTAAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTCAG 396
          || |||| || | | | ||| ||||| ||||| |||| ||| ||| | ||| || | ||
Db    276 GTTGCAGCAATGACCTTTGCCGGAAAGGCTTATTTTGATGCACTCTCAAAAATAGGGGAA 335

Qy    397 ATTGCCACTGGGGTCCCCCGTGTCAACTGAACTGGGACAATGTCCTCATAGAGATTTCAAGT 456
          ||||| |   |||| ||||| || | ||||  || |  | |||| | | |||| || ||
Db    336 AGTGCTGCTGTTTCGCCGGTTTCTAGGGAGCTGGGAGTGGTGTTAATGGAGATTTCAGA 395

Qy    457 ACCCACAAGAAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTCCACAAAGAGATT 516
          |||||||| | | |  ||  |||||||||| || | | |  |||  |||||  || || |
Db    396 AGTGCTGCTGTTTCGCCGGGTTTCTAGGGAGCTGGGAGTGGTGTTAATGGAGAA 395
```

```
Db    396  GTGCAACAAGAACGCCCAATTGGAGATTGGAGAGGAGTCTTTCACAAAAGAGCTC  455

Qy    517  ATCCATGAGCTGGAGAAGAGATAGAACTTGACGTGGAAATATGAACGCAACTCTAAAA  576
           ||||||     ||||||||||||| || |||||||||||||||||||||||||||||
Db    456  ATTGCTGAACTGGAAAAGAAAACTGATATATGGATACCAAGTACATGCCACTTTTAAA  515

Qy    577  AGATACCAAACAGAACACACAGAAATAAATTACGAGTGGAGAAATCCCAAGCTGAGTTG  636
           ||||||  ||||||||||||||||||||||     ||||||||||||||||||||||
Db    516  AGATACCAGTCAGAACACACAAATTAAAGCAAGACTTCCTGGACAAATCACAGGCTGACCTG  575

Qy    637  AAGAAGGATCAGAAGGAAAAAGCCAAGAAGCCACTCAAAATATGACACAAAGAA  696
           ||||        |||||||||||||||||||  ||||||||||||||||||||
Db    576  AAGAAAACTACGAAGAAAAAGTCAAGGGA---AACATTCGTCAAATATGAGATCAAAAGAA  632

Qy    697  ATTGAGTATGTGGAGACCGGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATTGCAGAT  756
           |||||      ||||||||||||||||  | ||||||||||||||||||| |||||||||
Db    633  AATGAGTGCATGGAGACCATCTCATCCCGGCAGACATCGAGACGTTTTATTGCAGAT  692

Qy    757  GGTTGCAAAGAGGGCTCTGCTTGAAGAGAAGAGGCGCTTCTGCTTTCTGGGTTGATTAAGCAC  816
           || ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    693  GGATGTPAAGGAGGGCTTTTACTAGAGAGAAAAGGAGGTTTGTTTCCTGGTGGAACAAACAC  752

Qy    817  TGTGGGCTTTGCAAACCACATACATTATTACTTACAGTCTGCAGAACTACTGAATTCC  876
           |||    ||||||||| |||||||| ||| |||||||||||||||||||||||||||
Db    753  TGTGCCCTTTCTTTATCATCAGCTTTCCATGACAAGGCCAAAGGAGGAATGCTGACCGGTC  812

Qy    877  AAGCTGCCTCGGTGGCCAGGAGGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATG  936
           ||||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db    813  AAACTGCCCAGCTGGCAAGAAAAGTGCAATGATGCCACCAGGGTGCCAGACACAGTGATA  872

Qy    937  AATATGATCGAAGAAAATAAAAGACCCCAGCCTCTACCCCGTCTGGGAACTCCTCAGGCT  996
           |||||||||||||||||               |||||||||||||||||||||||||||
Db    873  TCCATGATTGAGGGTCTGCGGGACTTCA-----ATGTCAGTGCATACCCAGAGTCT  920

Qy    997  TCACCCCATGATCGAGAAGCAATGTGGTTAGGAAAAGATTACGACACCCCTTTCTCAAATGC  1056
           |||        |||||||||| || ||  |||      | ||||||| ||||  || |||
Db    921  TCACCTACATTGGAACGCAGT-------------GACAGGACCAATTCCGGT  959

Qy    1057 TCACCAAAGATGCCCCCGCCTCCTTCAGGCAGAGCANATACCAGTCCCTTGATCGATATG  1116
           ||||||||||   |||||||||||||||||||| || |||||||||||||||||||||
Db    960  TCAATTGTTCCACCGCCACCCCCATCACTTAAAGCCCCACACGAGCCCAGCACATG  1019

Qy    1117 TTTAATAACCCACCACCGGCTGCCCGAATTCACAAAGGGTAAAATAATTCAAACAGGTACT  1176
           |||||  |||||||||||  |||||||||||||||||||||| || |||||| ||||||
Db    1020 TTCTCCCAGGAAGCTCCCAAAAGCCCATCACTTAAAGCCGAGCCCACGGCAGTTCAGATCAGGACAGT  1079

Qy    1177 TCCGAAGATCCCAGTTTACAGACGATCAGTTTCGGTTGCANACGGGACTGAACATGATGAAG  1236
           ||||        |||||||||||| ||||||| ||||||  ||||||||||| ||||||
Db    1080 CTAGACGAAAACGCCCTGTCTCGGTCTCAGTGTCTTAGCGCGCATGGAACACGGCGGAA  1139

Qy    1237 AAGCAGAAAAGTGAAGACCCATCTTCCCGGCAGCATCCTCCAACAAGACCTTACTCAGC  1296
           |||        |||||||||||||||||||||||  |  | |||||||||||||||||
Db    1140 AAGACCAGGGTGCAAACTATTTTTCCCCCACCACAGAATAACGAACGCTCGTGAGC  1199

Qy    1297 TTTGCACAGGGAGAGATGTCATCAGCGTCGTCATCCCCGAGGAGAAGGATGGCTCGGCTCTAT  1356
           ||||        |||||||||||||||||||||||| || | ||||||||||||||||
Db    1200 TTTGAGGATGGAGACATCATTATTTTGCTTTATCCAGGAAGAGCGGGACACGGCTGACTGTAC  1259

Qy    1357 GGAGAACACGACGTGTCCAAGGCCGAGGGGTTGGTTCCCGTCGTCGTA  1403
           ||||        || |||||||| |||| || |||||| ||||
Db    1260 GGAGAACTTGAAACATACCCGACAACGGGGGAATGGTTCCGTTCTTCTTA  1306
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: March 20, 2006, 13:44:26 ; Search time 2596 Seconds
        (without alignments)
        1882.610 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa..........aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
        Gapop 10.0 , Gapext 1.0
Searched: 8023312 seqs, 116585285A residues
Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
        Maximum Match 100%
        Listing first 45 summaries

Database :   Published_Applications_NA_New:*
    1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
    2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
    5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
    7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
    8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seqi:*
    9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seqi:*
    10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
    11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
    12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
    13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	357.4	17.1	516	12	US-11-021-492-430	Sequence 430, App
2	266.6	12.7	2078	12	US-11-128-061-1100	Sequence 1100, Ap
3	266.6	12.7	2078	12	US-11-128-049-1100	Sequence 1100, Ap
4	209.6	10.0	303	12	US-11-128-061-2130	Sequence 2130, Ap
5	209.6	10.0	303	12	US-11-128-061-5772	Sequence 5772, Ap
6	209.6	10.0	303	12	US-11-128-049-5772	Sequence 5772, Ap
7	209.6	10.0	303	12	US-11-128-049-2130	Sequence 2130, Ap
8	147.2	7.0	640	6	US-09-925-065A-730687	Sequence 730687,
c 9	81.4	3.9	1395	8	US-10-955-054A-111	Sequence 111, App
c 10	79.4	3.8	1400	12	US-11-128-061-4742	Sequence 4742, Ap
c 11	79.4	3.8	1400	12	US-11-128-049-4742	Sequence 4742, Ap
12	79.4	3.8	2036	8	US-10-996-217A-8	Sequence 8, Appli
13	78.8	3.8	1400	12	US-11-136-527-7203	Sequence 7203, Ap
14	78.8	3.8	1967	12	US-11-136-527-3107	Sequence 3107, Ap
15	78	3.7	2120	9	US-10-689-742-163	Sequence 163, App
16	77.8	3.7	932	9	US-11-031-206-65	Sequence 65, Appl
17	77.4	3.7	1133	8	US-11-663-794-1	Sequence 1, Appli
18	77.8	3.6	1400	12	US-11-136-527-6433	Sequence 6433, Ap
19	76	3.6	3801	12	US-11-136-527-2337	Sequence 2337, Ap
20	75.8	3.6	2255	9	US-11-096-568A-2260	Sequence 2260, Ap
21	75.6	3.6	2797	12	US-11-120-308-133	Sequence 133, App
22	75.6	3.6	4330	12	US-11-091-883-182	Sequence 182, App
23	75.4	3.6	1456	9	US-11-096-568A-12004	Sequence 12004, A
24	75.4	3.6	2091	9	US-10-276-233A-19	Sequence 19, Appl
25	75	3.6	1273	9	US-11-096-568A-12247	Sequence 12247, A
26	75	3.6	4339	8	US-10-909-125-801	Sequence 801, App
27	74.8	3.6	2162	12	US-11-091-883-330	Sequence 330, App
28	74.6	3.6	600	12	US-11-123-896-256	Sequence 256, App
29	74.6	3.6	755	7	US-10-973-115B-153	Sequence 153, App
30	74.6	3.6	755	12	US-10-131-826A-153	Sequence 153, App
31	74.6	3.6	3082	9	US-11-096-568A-2443	Sequence 2443, Ap
32	74.4	3.5	600	12	US-11-136-527-7965	Sequence 7965, Ap
33	74.4	3.5	625	9	US-11-096-568A-2599	Sequence 2599, Ap
34	74.4	3.5	804	12	US-11-091-883-265	Sequence 265, App
35	74.4	3.5	2183	12	US-11-000-688-485	Sequence 485, App
36	74.4	3.5	2924	12	US-10-973-115B-37	Sequence 37, Appl
37	74.4	3.5	3501	7	US-11-136-527-3869	Sequence 3869, Ap
38	74.4	3.5	3501	12	US-10-131-826A-37	Sequence 37, Appl
39	74	3.5	850	8	US-11-096-568A-3385	Sequence 3385, Ap
40	74	3.5	1358	9	US-11-096-568A-2642	Sequence 2642, Ap
41	74	3.5	2537	12	US-11-102-978-8	Sequence 8, Appli
42	73.8	3.5	799	12	US-11-195-459-3	Sequence 3, Appli
c 43	73.8	3.5	807	9	US-11-096-568A-706	Sequence 706, App
c 44	73.8	3.5	1660	9	US-11-096-568A-2748	Sequence 2748, Ap
45	73.8	3.5	2476	7	US-10-973-115B-489	Sequence 489, App

ALIGNMENTS

RESULT 1
US-11-021-492-430

    Sequence 430, Application US/11021492.

; Publication No. US20060031947A1
; GENERAL INFORMATION:
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0368-USA
; CURRENT APPLICATION NUMBER: US/11/021,492
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/307,670
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 698
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40, 156, 175, 190, 216, 291, 344, 370, 431
; OTHER INFORMATION: n = A,T,C or G
US-11-021-492-430

Query Match   83.9%;   Score 357.4;   DB 9;   Length 516;
Best Local Similarity   17.1%;
Matches   433;   Conservative   0;   Mismatches   80;   Indels   3;   Gaps   3;

```
Qy    429 GGGACAATGGTCCTCAATAAGAGGATTTCAAGTACCCAACAAAGAAACTCAACGAAGAGTCTTGATGA 488
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGGCCTGTCCTCAATAAGAGATTTCCAGCACCCAAGGAANCTCAATGAGACTCTTGATGA 60

Qy    489 AAATTTAAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAAGAATAGAAGACTTGA 548
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAATTTCAAAAAAATTCCATAAGGATATTATCCATGAAGCTGGAGAAGAAGAGAACTTGA 120

Qy    549 CGTGAAAATATATGAA-CGCAACTCTTCAAAAAGATACCAAACAGACAACAAGAATAAAATT-A 606
          ||  |||||||||||| ||  ||||||||||||||||||||  ||||||||||||||||||| |
Db    121 TGTAAAGTATATATGAACCGCTACTCTTAAAAAGATATNAAGCGAGAGCACAGGAATANNATTAA 180
```

Left column

```
Qy  607 GAGTCTTTGGAGAGAAATCCCAAGC-TGAGTTGAAGAAGATCAGAAGGAAAAGCAAGGAAG 665
        |||||||| |||||||||| ||| ||  |||||||||| |||||||||| | || |||| ||
Db  181 GATTCTTTGNAGAAATCTTCAGGCAGGAGCTGAAGANGAATCAGAGANGAAGAAAAGTGG 240

Qy  666 CCGAAACGCACTCAAATATGAACACAAGAAAATTGAGTAGTGGAGACCGTTACTTCTCG 725
        ||||| ||||||||||||| || ||| || |||||||||| |||||||| ||||  |||||
Db  241 ACGAAAACGCACTCAAATACGACGACAAAGAAAATTGAGTATGTGGAGACGGGNGACGTCTCG 300

Qy  726 TCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGAGGCTCTGCCTTGAAGAGAA 785
        |||||||||| |||||||| |||||||| ||||| ||||||||| ||| ||||||||| || ||
Db  301 CCAGAGTGAGATCCAGAAGTTCATTGCCGATGGCTGCAAAGAACCCTGCTTGAAGGAAA 360

Qy  786 GAGGCGCTTCTGCCTTTCTTGGTTGTTGCAAACCACATACATTATTA 845
        ||||||||| | |||| ||||||| | ||||| |||| ||||| | ||
Db  361 AAGGCGCTTTNTGCTTTCTTGGTGGACAAGCACTGCACATACTACTA 420

Qy  846 TCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGAGGAGACCTGTGT 905
        |||| | ||| ||||||| |||| | |||||||||||||| ||||||| ||||| ||| |
Db  421 TCACATGCAGNCTGCAGAGTTACTTAATTCCAAGCTGCCCCGGTGGCAGGAAAACCTGTTG 480

Qy  906 TGATGCCATCAAAGTGCCAGAGAGAAAAATGAATTAT 941
        |||||||| | ||| |||||||||||||| |||| ||
Db  481 TGATGCCACAAAAAGTCCCAGAGAGAAAATCATGAACAT 516
```

RESULT 2
US-11-128-061-1100
; Sequence 1100, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1100
; LENGTH: 2078
; TYPE: DNA
; ORGANISM: Cricetinae
US-11-128-061-1100

Query Match   12.7%; Score 266.6; DB 12; Length 2078;
Best Local Similarity   59.2%; Pred. No. 5.3e-24;
Matches 455; Conservative   0; Mismatches 314; Indels   0; Gaps   0;

```
Qy  230 CCGAGGAGGTGAACCGGCTCACCGGAGAGCACCGTTACTTCAATC 289
        ||||||||||  ||||||||||| || |||   |    |||||||||
Db   77 CCGGAGGAGGATGCACCGGCTCAGGGGAAAATGTCTATATAAGACCATCATGGAGAACAGTTCAACC 136

Qy  290 CTGGGGCTGCGAAATTTAATAAACCTGGGGAAAATGTTCTATATAAGGCTATGA 349
Db  137 CCAGCCTCCGCAACTTCATCGCCATGGGGAAGAACTACGAGAAAGCCCTCCGTCAGGTGTCA 196

Qy  350 TCCTGGCCGGAAAAAGCCTACTACGGATGGAGTGGCCAGAGATCGCGTCGATGGTTGCCACTCGGT 409
Db  197 CCTTCGCTGCCAAAAGGCTATTCGATGCTCTGTAAAGATGGGGAGGAGCTGGGCCGCCGAGA 256

Qy  410 CCCCGGTGTCAACTGAACTGGGGACAGTGTCCTCATAGAGAGATTTCAAGTAGTACCCACAAGAAAC 469
Db  257 GCCAGGGGCTCTAAGGAACTTGGGGACGTTCTTCCTTCCACAGATGGGCCGAAGTCCACCGGGCGAGA 316
```

Right column

```
Qy  470 TCAACGACAGTCTTTGATGAGAAAAAATTTTAAAAAAATTCCACAGAGAGATTACCATGAGCTGG 529
Db  317 TCCAGAACCACTGGAAAGAGAGATGCGTAAAAGTCTTTTCACAAATGAGCTGTTCACACAGTCGG 376

Qy  530 AGAAGAAGATAGGAAGCTTGAGCGTGAAAATATATGAACGCAACTCTAAAAAAGATACCAGGCAG 589
Db  377 AACAGAAGGTAGACGCTGACTCCAGGATAGTCTGAGTGCTGCAAGGAATACCAGGCAG 436

Qy  590 AACAAGAGAATAAAATTAGGAGTCTTTGGAGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
Db  437 AGCAGGAGGAGCAAGAGGAAGATGCTCTGGAGGAACAAGAGTGTCAGGCTGAGGCTCGGTA 496

Qy  650 GGAAAAGCCAAGGAAGCCGAAAACGCACTCAAATATGAACACAAGAAAATTGAGTAGTGGATATGTGG 709
Db  497 AGAAGAGCCGAAGGGAGCAAGAAACCTCAGAAGTGACTCGGAGAACTACTGGAGTACATTG 556

Qy  710 AGACCGTTACTTCTCGCAGAGTGAAATCCAGAAATTCATTGCAGATGGTTGCAAAGAGAGG 769
Db  557 ATGCCATCAGCAACAGAGGGGTGAGCGTGGAACTACCGTATCTGATAGGCTACAAGACAG 616

Qy  770 CTCTGCCTTGAAGAGAAGAGGCGCTTCTGCCTTTCTTGGTTGATAAGCACTGTTGGCTTTGCAA 829
Db  617 CACTCACCGAGGAGGGCGGAGGAAGAGGGCTTCGCTTCCTTGGGAAAAAGCAGTGTGCTGCTGTGGCCA 676

Qy  830 ACCACATACATTATTACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGT 889
Db  677 AGAACTCTGCGCTTACCACTCCAAAGGGCAAGGAGCTGGTAGCCCAGAAGAGCTGCCGTGT 736

Qy  890 GGGAGGAGACCCTGTTGTGATGCCATCAAAGTGCCAGAGAGAAAATCATGAATATATATATATCGAG 949
Db  737 GGCAGCAACCCTGTTGCTGACCCAACAACAGAAATTCCAGACCGCGTGTCCAGCTGATGCAGC 796

Qy  950 AAAATAAAGACCCAGCCTCTACCCCCGTGTCTGGAAACTCCTCCAGGGCTTC 998
Db  797 AGATAGCCAGCACAACGGACGCTCCAATCCTTCCTAGTACCCTGTCAGCCTC 845
```

RESULT 3
US-11-128-049-1100
; Sequence 1100, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS F
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1100
; LENGTH: 2078
; TYPE: DNA
; ORGANISM: Cricetinae
US-11-128-049-1100

Query Match   12.7%; Score 266.6; DB 12; Length 2078;
Best Local Similarity   59.2%; Pred. No. 5.3e-24;
Matches 455; Conservative   0; Mismatches 314; Indels   0; Gaps   0;

```
Qy  230 CCGAGGAGGTGAACCGGCTCACCGGAGAGCACCGTTACTTCAATC 289
        ||||||||||  ||||||||||| || |||
Db   77 CCGGAGGAGGATGCACCGGCTCAGGGGAAAATGTCTATATAAGACCATCATGGAGAACAGTTCAACC 136
```

Left column (RESULT 4 alignment)

```
Qy   290  CTGGGCTGCGAAATTAATAAACCTGGGGAGAAAAATTATGAGAAGCTGTAAACGCTATGA  349
          |||||||||||||||||||||||
Db   137  CCAGCCTCCGCAACTTCATCGCCATGGGGAGAAGAACTACGAGAAAACCCTGGCAGTGTCA  196

Qy   350  TCCTGGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGT  409
Db   197  CCTTCGCTGCCAAAGGCTATTCGATGCTCTGGGTAAAGATGGGGGGAGCTGGCCACGCGAGA  256

Qy   410  CCCCCGTGTCAACTTAAACCTGGACATGTCTCCATAGAGATTTCAAGTACCCACAAGAAAC  469
Db   257  GCCAGGGCTCTTAAGGAACTTGGGGACGTCCTTCTTCCAGATGGCCGAAGTCCACCGGCAGA  316

Qy   470  TCAACGAGAGTCTTGATGAAAATTTTAAAAAATTTCACAAAAGAGGATTATCCATGAGCTGG  529
Db   317  TCCAGAAACAACTGGAAGAGATGCTAAAGTCTTTTTCACAAATGAGCTGCTCACACAGCTGG  376

Qy   530  AGAAGAGATAGAGAACTTGACGTGAAATATATGAACGGAACTCTAAAAGATACCAAACAG  589
Db   377  AACAGAAAGGTAGAGCTGGACTCCAGGTATCTGCACTGAAGAAGAAATACCAGGCAG  436

Qy   590  AACACAAGAAATAAAATTGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAGAAGATCAGAA  649
Db   437  AGCAGGAGGAGCAAAGGAGATGCTCTGGACAAAGTCAGCGCTGAAGAAGCTTCCGGTA  496

Qy   650  GGAAAAGCCAAGGAAGCCGACACTCAAATATGACAACAAAGGAAATTGAGTATGTGG  709
Db   497  AGAAGAGCCAAGGGAGACAGAAACCCTCAGAAGTACTCGGACAACAAGGAGCTGCAGTACATTG  556

Qy   710  AGACCGGTTACTTCTGTCAGAGTGAAATCCAGAAATTCATTGCACAGATGGTTGCAAAGAGG  769
Db   557  ATGCCATCAGCACAACAGCAGGGTGAGCTGGAGAACTACGTATCTGATGGCCTACAAGACAG  616

Qy   770  CTCTGCTTGAAGAGAAGAGAGCGCGTTCTTGCTTTCGGTTTCTGGTTTGATAAGCACTGTG  829
Db   617  CACTCACCGGAGGAGGAGGTTCTGCTTCCTCGGTTGGAAAAGGCAGTGTGCGTGGCCA  676

Qy   830  ACCACATACATTATTATCACTTACGTCTCAGAGAACTACTGAATTCCAAGCTGCTTCGGT  889
Db   677  AGAACTCTGCTGCTTACCACTCCAAAGGCAAGGAGCTGTTAGCCCCAGAAGCTGCCGTGTT  736

Qy   890  GGCAGGAGGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATCGAAG  949
Db   737  GGCAGCAAGCCTGTGGCTGACCCAACAACAGATTCCAGACGCGGCGTGTCCAGCTGATGCAG  796

Qy   950  AAATAAAGACCCCAGCCTCTACCCCGTGGTGTCTGGAACTCCTCAGGCCTTC  998
Db   797  AGATAGCCAGCACGACAACGGCTCCATCCTTCCTAGTACCCTGTCAGCCTC  845
```

RESULT 4
US-11-128-061-2130
Sequence 2130, Application US/11128061
Publication No. US20060003958A1
GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 0197.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2130
; LENGTH: 303

Right column

; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-2130

Query Match   10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity   83.1%; Pred. No. 4.3e-17;
Matches   250; Conservative   0; Mismatches   50; Indels   1; Gaps   1;

```
Qy   674  CACTCAAATATGAACACAAGAAAGAAAAATTGAGTATGTGGAGACCGGTTACTTCTGTCAGAGTG  733
Db     1  CACTGAAATATGAGCACACAAAGAAAATTGAGTATGTTGGAGACTGTTACTTCTCGGCCACAGAGTG  60

Qy   734  AAATCCAGAAAATTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGGGCT  793
Db    61  AAATCCATAAGTTCATTGCAGATGCGCTTGGCAAAGGAAGCCCCTGCTTGAAGGA-AAAAGCGCT  119

Qy   794  TCTTGCTTTCGGTTTGATTAAGCACTGTGGCTTTTGCAAACCACATACATTATTATCACTTAC  853
Db   120  TCTGCTTTCTGGTTGGACAAGCACTGTGGCTTTGCCAATCACATACTACTAATCACCTTC  179

Qy   854  AGTCTGCAGAACTACTGAATTCCAAGCTGCCTTCGGTTGGCAGGAGGACCTGTGTTGATGCCA  913
Db   180  AGTCTGCNAAATTACTTAATTCCAACCTGTCCCGGAGCGGCAGAAAACCTGTTGATGCCA  239

Qy   914  TCAAAGTGCCAGAGAAAATCATGAATAATGATCGAAGAAATAAAGACCCCAGCCTCTACCC  973
Db   240  CATAAGCACCAGAGAAAATCATGAACAATGATAGAGACGCCCATCTTCACAC  299

Qy   974  C  974
          |
Db   300  C  300
```

RESULT 5
US-11-128-061-5772
Sequence 5772, Application US/11128061
Publication No. US20060003958A1
GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 0197.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5772
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-5772

Query Match   10.0%; Score 209.6; DB 12; Length 303;
Best Local Similarity   83.1%; Pred No. 4.3e-17;
Matches   250; Conservative   0; Mismatches   50; Indels   1; Gaps   1;

```
Qy   674 CACTCAAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTTCGTCAGAGTG 733
         ||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db     1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTTCGCCAGAGTG 60

Qy   734 AAATCCAGAAATTCATTGCGAGATGGTTGCAAAGAGGGCTCTGCTTGAAGAGAGAGGCGCT 793
         |||||| || |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    61 AAATCCATAAGTTCATTGCGAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAGCGCT 119

Qy   794 TCTGCTTTCTGGTTGATAAGCACTGTGGGCTTTGCAAACCACATACATTATTATCACTTAC 853
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   120 TCTGCTTTCTGGTTGGTGGACAAGCACTGTAGCTTTTGCCAATCACTACTACTATCACTTC 179

Qy   854 AGTCTGCAGAACTACTGAATTCCAAGCTCCCTCGGTGGCAGGAGACCTTGTTGATGCCA 913
         |||||||| |||| || ||||||||||||||| |||||||||||||||||||||| ||||
Db   180 AGTCTGCNAAAATTACTTAATTCCAACCTGTCCCGGAGGAGAAACCTGTTCGAGATGCCA 239

Qy   914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy   974 C 974
         |
Db   300 C 300
```

Query Match   10.0%;   Score 209.6;   DB 12;   Length 303;
Best Local Similarity   83.1%;   Pred. No. 4.3e-17;
Matches 250;   Conservative   0;   Mismatches 50;   Indels   1;   Gaps   1;

```
Qy   674 CACTCAAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTTCGTCAGAGTG 733
         ||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db     1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTTCGCCAGAGTG 60

Qy   734 AAATCCAGAAATTCATTGCGAGATGGTTGCAAAGAGGGCTCTGCTTGAAGAGAGAGGCGCT 793
         |||||| || |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    61 AAATCCATAAGTTCATTGCGAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAGCGCT 119

Qy   794 TCTGCTTTCTGGTTGATAAGCACTGTGGGCTTTGCAAACCACATACATTATTATCACTTAC 853
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   120 TCTGCTTTCTGGTTGGTGGACAAGCACTGTAGCTTTTGCCAATCACTACTACTATCACTTC 179

Qy   854 AGTCTGCAGAACTACTGAATTCCAAGCTCCCTCGGTGGCAGGAGACCTTGTTGATGCCA 913
         |||||||| |||| || ||||||||||||||| |||||||||||||||||||||| ||||
Db   180 AGTCTGCNAAAATTACTTAATTCCAACCTGTCCCGGAGGAGAAACCTGTTCGAGATGCCA 239

Qy   914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy   974 C 974
         |
Db   300 C 300
```

Query Match   10.0%;   Score 209.6;   DB 12;   Length 303;
Best Local Similarity   83.1%;   Pred. No. 4.3e-17;
Matches 250;   Conservative   0;   Mismatches 50;   Indels   1;   Gaps   1;

```
Qy   674 CACTCAAAATATGAACACAAAGAAATTGAGTATGTGGAGACCGTTACTTCTTCGTCAGAGTG 733
         ||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||||
Db     1 CACTGAAATATGAGCACAAAGAAATTGAGTATGTGGAGACTGTTACTTCTTCGCCAGAGTG 60

Qy   734 AAATCCAGAAATTCATTGCGAGATGGTTGCAAAGAGGGCTCTGCTTGAAGAGAGAGGCGCT 793
         |||||| || |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    61 AAATCCATAAGTTCATTGCGAGATGGCTGCAAAGAAGCCCTGCTTGAGGA-AAAAGCGCT 119

Qy   794 TCTGCTTTCTGGTTGATAAGCACTGTGGGCTTTGCAAACCACATACATTATTATCACTTAC 853
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   120 TCTGCTTTCTGGTTGGTGGACAAGCACTGTAGCTTTTGCCAATCACTACTACTATCACTTC 179

Qy   854 AGTCTGCAGAACTACTGAATTCCAAGCTCCCTCGGTGGCAGGAGACCTTGTTGATGCCA 913
         |||||||| |||| || ||||||||||||||| |||||||||||||||||||||| ||||
Db   180 AGTCTGCNAAAATTACTTAATTCCAACCTGTCCCGGAGGAGAAACCTGTTCGAGATGCCA 239

Qy   914 TCAAAGTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCC 973
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   240 CATAAGCACCAGAGAAAATCATGAACATGATAGAAGCGATAAAAACGCCCATCTTCACAC 299

Qy   974 C 974
         |
Db   300 C 300
```

RESULT 8
US-09-925-065A-730687/c
; Sequence 730687, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 730687
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-730687

Query Match   7.0%; Score 147.2; DB 6; Length 640;
Best Local Similarity   92.3%;
Matches 155; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy   709 GAGACCGTTACTTCTCGTCAGAGTGAAAATCCAGAAAATTCATTGCAGATACGGTTGCAAAGAG 768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   640 GAGACCGTTACTTCTCGTCAGAGTGAAAATCCAGAAAATTCATTGCAGATGGTTGCAAAGAG 581

Qy   769 GCTCTGCTTGAAGAGAGAAGAGGCCGCTTCTCGCTTCTTTTGGTTGATAAGCACTGTGGCTTTGCA 828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   580 GCTCTGCTTGAAGAGAGAAGAGGCCGCTTCTCGCTTCTTTTGGTTGATAAGCACTGTGGCTTTGCA 521

Qy   829 AACCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 876
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   520 AACCACATACATTATTATCACTTACAGGTGGGGTGTGGCAAAATTGC 473

Qy   1942 ATGAGTRAATAGTCTAATGACCAGCTCAGCCATTAAATATTTTCTTCCTATTCTGTTCA 2001
          |||||||||| |||||
Db   1216 AGGAATAGCAGAATAATTGGTGGTCCTTGTAAACATTTAAGATTCTCCGTGGATTTGGTGA 1275

Qy   2002 AGAAACAGTAAACTGGTTTCAAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2061
          ||||| |||||||||||||||||||||||||||||||
Db   1276 GTGATCATTAAACTGTTTTCCAACTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1335

Qy   2062 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
          |||||||||||||||||||||||||||||||||||||||||
Db   1336 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1370

RESULT 9
US-10-955-054A-111
; Sequence 111, Application US/10955054A
; Publication No. US20050266420A1
; GENERAL INFORMATION:
; APPLICANT: PUSZTAI, LAJOS
; APPLICANT: SYMMANS, W. FRASER
; APPLICANT: HESS, KENNETH R.
; APPLICANT: AYERS, MARK
; APPLICANT: STEC, JAMES
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
; FILE REFERENCE: UTXC:880US
; CURRENT APPLICATION NUMBER: US/10/955,054A
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 1395
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-955-054A-111

Query Match   3.9%; Score 81.4; DB 8; Length 1395;
Best Local Similarity   70.3%;
Matches 109; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

RESULT 10
US-11-128-061-4742
; Sequence 4742, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4742
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Cricetinae
US-11-128-061-4742

Query Match   3.8%; Score 79.4; DB 12; Length 1400;
Best Local Similarity   54.5%;
Matches 186; Conservative 0; Mismatches 146; Indels 9; Gaps 1;

Qy   1193 TACACGCATGATCAGTTTCGGTTGCAACGGACGTGACACATGATGAAGAAGCAGAAAGTCAAGA 1252
          ||||||||| ||||| ||||||| ||||||||||||||||||||
Db   467 TGCCACGTTCAAGTCCCATGCAGCTGGCCTTAGAACGTAACGGCCGGTCAAAG 526

Qy   1253 CCAICTTCCCGCACACTGCGGGGGCTCCAACAAAGACCTTACTCAGCTTTGCACAGGGATG 1312
          ||||| |||||||||||||||||||||| ||||
Db   527 CCATTTTCTCCCACCCGGCGTGGTGACAACAGCACTCTGCTGAGCTTCAACGAGGCGGACC 586

Qy   1313 TCATCACGCTGCTCATCCCCGAGGAGAAGGATGGCTGGGCTTCTATGGAGAACACGCGTGT 1372
          ||||||||||||||||| |||||||||||||||||| || |||||||||||||||
Db   587 TCATCACCCTGCTCAGTGCCTGAGCGCCCGTGACGGCGGCCACTAT GGCGAGGAGTGGAAAA 646

Qy   1373 CCAAGGCGAGGGGTTGGTTCCCGTGCGTCGTGTACACGGAAGTTGCTGGAAGAAAATGAGACAG 1432
          ||||||||||||||| || || |||| ||||||||||||||||||
Db   647 CCAAAATGCGGGGCGTGGTTTCCTTTCCTACACTGCGGTCTGGCACAGTGACGGAAGTG 706

Qy   1433 AAGCAGTGACGACCGTGCCCACGCCCAAGCCCCACACCAGTGAGAAGGCATCAGCCACCGTGAACT 1492
          ||||||||||||||||| ||||||||||||||||||||||||||||||
Db   707 ACAGGTTGCACATG---------AGCCTCCAGCAGGGCAGGAGCAGCCACGGCAGCGGCCGCAAC 757

Qy   1493 TGTCTGAGAATAGCAGTGGTTGTTGTTCATCCCCCCACCCGACTAC 1533
          ||||||||||||||||||||||| ||||||||||||||||||
Db   758 TTCTGGACAAGGAGCGACCTGGCCGTCCCCCCTCCCCTGACTAC 798

RESULT 11
US-11-128-049-4742
; Sequence 4742, Application US/11128049

; Publication No. US20060010511A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 4742
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Cricetinae
US-11-128-049-4742

Query Match   3.8%;   Score 79.4; DB 12; Length 1400;
Best Local Similarity 54.5%;   Pred. No. 0.17;
Matches 186; Conservative 0; Mismatches 146; Indels 9; Gaps 1;

Qy   1193 TACAGCGATCAGTTTCGGTTGCAACGGGACTGAACACATGATGAAGAAGCAGAAAGTGAGA 1252
               ||| || || || || || | ||| |||| ||| || || |||| |||| ||
Db   467 TGCCACGTTCAAGCTCCATTGGCAGCTGGCCTAGAACGTAACGCGCCGCATGCGGCTCAAAG 526

Qy   1253 CCATCTTCCCGCACACTGCGGGGGTTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATG 1312
               |||||| | ||| | |||| || | | ||| ||| ||| || |||
Db   527 CCATTTTCTCCCACGCGGCTGGTGATGACAACAGCACTCTTCGTGAGCTTCAAGGAGAGGCGACC 586

Qy   1313 TCAATCACGCTGCTCATCCCGAGAGAAGGAGATGGCTGGCTTCCTATCGGAGAACACGACGTGT 1372
               || || |||| ||| ||| || |||| |||| || ||||| ||| ||||
Db   587 TCATCACCCTGCTAGTGCCTGAGGCCCGTGACGCGCTATCGGCGAAGCTGGAGTGAGAAA 646

Qy   1373 CCAAGGCGGAGGGGTTGGGTTCCCGTCGTGTGCTACACGGAAGTTGCTGGAGAAAAATGAGACAG 1432
               || ||| | || |||| |||| || ||| ||| ||| ||| || ||| |||
Db   647 CCAAAATGCGGGCGGCTGGTTTCCTTCTCCTACACTCGGGTCCTGGACAGTGCACGGAAGTG 706

Qy   1433 AAGCAGTGACGTGCCCACGCCCACACCAGTGAGAAGCATCAGCACCCGTGAACT 1492
               || ||| ||| || ||||| | |||| ||| |||| |||| ||||| |||||
Db   707 ACAGGTTGCACATG------AGCCTCCAGCAGGGCAAGAGAGCAGCACGGGGCCAACC 757

Qy   1493 TGTCTGAGAATAGCAGTGTTGTCATCCCCCACCGGACTAC 1533
               | |||| | ||| |||||||| |||||||||||||||
Db   758 TTCTGGACAAGGACGACCTGGCCGGTCCCCCCTCCTGACTAC 798

; RESULT 12
US-10-996-217A-8/c
; Sequence 8, Application US/10996217A
; Publication No. US20050266561A1
; GENERAL INFORMATION:
; APPLICANT: Revivicor, Inc.
; APPLICANT: Wells, Kevin
; TITLE OF INVENTION: Use of Interfering RNA in the Production of Transgenic Animals
; FILE REFERENCE: 10785.105070 REV 1015 US
; CURRENT APPLICATION NUMBER: US/10/996,217A
; CURRENT FILING DATE: 2004-11-22
; PRIOR APPLICATION NUMBER: 60/523,938
; PRIOR FILING DATE: 2003-11-21
; NUMBER OF SEQ ID NOS: 272
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-996-217A-8

Query Match   3.8%;   Score 79.4; DB 8; Length 2036;
Best Local Similarity 73.7%;   Pred. No. 0.16;
Matches 101; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy   1960 ACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAAACTTGGT 2019
               |||| || |||| | ||||||||||||| ||| || || ||| |||||| |||| |||
Db   668 ACCATTACCCATTTTAGGATATTTTTCTTGGTTCCCTTCTCCAGTGAAGAAACTTATTTTAG 609

Qy   2020 TTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2079
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GTAATCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 549

Qy   2080 AAAAAAAAAAAAAAAAAA 2096
               ||||||||||||||||||
Db   548 AAAAAAAAAAAAAAAAAA 532

; RESULT 13
US-11-136-527-7203
; Sequence 7203, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 7203
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7203

Query Match   3.8%;   Score 78.8; DB 12; Length 1400;
Best Local Similarity 65.1%;   Pred. No. 0.2;
Matches 82; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

Qy   1971 CATTTAAAATATTTTCTTCCTATTTCTGTTCAAGAAACAGTAAAACTTGGTTTCAATCTTTA 2030
               || ::: : : :|| ::::
Db   1254 CTTTTGAWACATGKACTTTCTGWTTTWYTCAARTAAAAGTTGGAAACTAYWRMWHVWMMW 1313

Qy   2031 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2090
               :::::|:|: :|||||||||||||||||||||||||||||||||||||||||||||
Db   1314 MWMWMWMWMRAWWAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1373

Qy   2091 AAAAAA 2096
               ||||||
Db   1374 AAAAAA 1379

; RESULT 14
US-11-136-527-3107
; Sequence 3107, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830

```
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3107
; LENGTH: 1967
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3107

Query Match        3.8%;    Score 78.8;  DB 12;  Length 1967;
Best Local Similarity  65.1%;   Pred. No. 0.19;
Matches  82;  Conservative  20;  Mismatches  24;  Indels  0;  Gaps  0;

Qy  1971 CATTTAAAATATTTTCTTCCTATTCTGTTCCAAGAAACAGTAAAACTTGGTTTCAATCTTTA 2030
              ||||||:|||::|:.|:|.:|:|.:|||:|||.:||:|:||||||  :::||  ::::  :
Db  1821 CTTTTGAWACATGKACTTTCTGWTTTWYTCAARTAAAAGTTGGAAACTAYWRMWHVMMMM 1880

Qy  2031 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2090
              :::::::|:|||:|||||||||||||||||||||||||||||||||||||||||||
Db  1881 MWWMWMARAWAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1940

Qy  2091 AAAAAA 2096
              ||||||
Db  1941 AAAAAA 1946

RESULT 15
US-10-689-742-163
Sequence 163, Application US/10689742
Publication No. US20050250180A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M
APPLICANT: LaVallie, Edward R
APPLICANT: Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 00766.000091.10
CURRENT APPLICATION NUMBER: US/10/689,742
CURRENT FILING DATE: 2003-10-22
PRIOR APPLICATION NUMBER: 09/746,783
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patentin version 3.2
SEQ ID NO 163
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
US-10-689-742-163

Query Match        3.7%;    Score 78;  DB 8;  Length 2120;
Best Local Similarity  59.5%;   Pred. No. 0.24;
Matches  132;  Conservative  0;  Mismatches  90;  Indels  0;  Gaps  0;

Qy  1875 CTGCTTAAGCAAATCATGCTTCTTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTT 1934
              ||||:|:::||:|:||.|:||||:|::||:||.|:|:|.|.|.:||..||:||:||||
Db  1887 CTATGTAAATATAAAGACTACTGGCTTTGTGAGGGAATGTTTGTGCAAAATTTTTTCCTCT 1946

Qy  1935 AAGATAAAATGAGTAAATAGTCTAAATGACCAGCTCAGCCATTTAAAATATTTTCTTCCTATT 1994
              |||:||||  |||| ||||:|||||:|||.||||||:||||||||||||:|||||||||
Db  1947 AATGTATAATAGTGTTAAAATGATTAAAAATCTTCAGAATTAAATTAATTCCCTTTGTCAC 2006

Qy  1995 CTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAA 2054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2007 TTTTTGAAAACATAATAAATCATTTGTATCTGTGCCTTAAAAAAAAAAAAAAAAAAAAA 2066

Qy  2055 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||
Db  2067 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2108
```

Search completed: March 20, 2006, 14:27:56

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2006 Bioacceleration Ltd.

GenCore version 5.1.7

Run on: March 20, 2006, 13:39:55 ; Search time 381 Seconds
                                              (without alignments)
                                              9778.922 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa..........aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY NUC
                    Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                    1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
                    2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
                    3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                    4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                    5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
                    6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                    7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
                    8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
                    9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	% Match	Length	DB	ID	Description
1	259.6	12.4	12.4	2080	2	US-08-878-563A-2	Sequence 2, Appli
2	259.6	12.4	12.4	2080	3	US-09-270-117-2	Sequence 2, Appli
3	259.6	12.4	12.4	2080	3	US-08-821-994-64	Sequence 64, Appl
4	241.6	11.5	11.5	2328	3	US-09-046-572-2	Sequence 4, Appli
5	82.8	4.0	4.0	997	3	US-09-907-794A-376	Sequence 376, App
6	82.8	4.0	4.0	997	3	US-09-905-125A-376	Sequence 376, App
7	82.8	4.0	4.0	997	3	US-09-902-775A-376	Sequence 376, App
8	82.8	4.0	4.0	997	3	US-09-906-700-376	Sequence 376, App
9	82.8	4.0	4.0	997	3	US-09-903-603A-376	Sequence 376, App
10	82.8	4.0	4.0	997	3	US-09-904-920A-376	Sequence 376, App
11	82.8	4.0	4.0	997	3	US-09-909-064-376	Sequence 376, App
12	82.8	4.0	4.0	997	3	US-09-905-381A-376	Sequence 376, App
13	82.8	4.0	4.0	997	3	US-09-906-618-376	Sequence 376, App
14	82.8	4.0	4.0	997	3	US-09-906-646-376	Sequence 376, App
15	82.8	4.0	4.0	997	3	US-09-904-462-376	Sequence 376, App
16	82.8	4.0	4.0	997	3	US-09-902-736A-376	Sequence 376, App
17	82.8	4.0	4.0	997	3	US-09-906-722A-376	Sequence 376, App
18	81.2	3.9	3.9	1570	3	US-10-012-231A-291	Sequence 291, App
19	81.2	3.9	3.9	1570	3	US-10-015-389A-291	Sequence 291, App
20	81.2	3.9	3.9	1570	3	US-10-006-768A-291	Sequence 291, App
21	81.2	3.9	3.9	1570	3	US-10-015-671A-291	Sequence 291, App
22	81.2	3.9	3.9	1570	3	US-10-015-393A-291	Sequence 291, App
23	81.2	3.9	3.9	1570	3	US-10-011-833A-291	Sequence 291, App
24	81.2	3.9	3.9	1570	3	US-10-006-041A-291	Sequence 291, App
25	81.2	3.9	3.9	1570	3	US-10-012-064A-291	Sequence 291, App
26	80.8	3.9	3.9	687	3	US-09-907-907A-38	Sequence 38, Appl
27	80.2	3.8	3.8	1474	3	US-08-821-994-64	Sequence 64, Appl
28	79	3.8	3.8	2550	9	5258287-23	Patent No. 5258287
29	78.8	3.8	3.8	1172	2	US-07-945-288-9	Sequence 9, Appli
30	78.8	3.8	3.8	1172	2	US-08-462-831-9	Sequence 9, Appli
31	78.8	3.8	3.8	1172	2	US-08-461-809-9	Sequence 9, Appli
32	78.8	3.8	3.8	1172	6	PCT-US93-08518-9	Sequence 9, Appli
33	78.8	3.8	3.8	1172	2	US-08-461-441-9	Sequence 9, Appli
34	78.6	3.8	3.8	396	3	US-09-970-966-53	Sequence 53, Appl
35	78.6	3.8	3.8	396	3	US-09-640-173-10	Sequence 10, Appl
36	78.6	3.8	3.8	396	3	US-09-713-550-53	Sequence 53, Appl
37	78.6	3.8	3.8	396	3	US-09-713-550-10	Sequence 10, Appl
38	78.6	3.8	3.8	396	3	US-09-825-294-10	Sequence 10, Appl
39	78.6	3.8	3.8	396	3	US-09-825-294-53	Sequence 53, Appl
40	78.6	3.8	3.8	396	3	US-09-970-966-10	Sequence 10, Appl
41	78.6	3.8	3.8	396	3	US-09-970-966-53	Sequence 53, Appl
42	78.6	3.8	3.8	708	3	US-09-270-767-13081	Sequence 13081, A
43	78.2	3.7	3.7	985	3	US-09-832-129-19	Sequence 19, Appl
44	78.2	3.7	3.7	985	3	US-09-832-129-32	Sequence 32, Appl
45	78	3.7	3.7	2276	3	US-09-205-258-183	Sequence 183, App

ALIGNMENTS

RESULT 1
US-08-878-563A-2
; Sequence 2, Application US/0887856

; Patent No. 5891674

; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Lal, Preeti
;   APPLICANT: Shah, Purvi
;   TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/878,563A
;   FILING DATE: Filed Herewith
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0323 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2080 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: BRSTNOT04
;   CLONE: 918158
US-08-878-563A-2

Query Match   12.4%; Score 259.6; DB 2; Length 2080;

```
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/270,117
  FILING DATE:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/878,563
  FILING DATE:
ATTORNEY/AGENT INFORMATION:
  NAME: Billings, Lucy J.
  REGISTRATION NUMBER: 36,749
  REFERENCE/DOCKET NUMBER: PF-0323 US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-855-0555
  TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 2080 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  IMMEDIATE SOURCE:
  LIBRARY: BRSTNOT04
  CLONE: 918158
US-09-270-117-2
```

```
Query Match     12.4%;  Score 259.6; DB 3;  Length 2080;
Best Local Similarity   59.2%;  Pred. No. 1.6e-48;
Matches  442;  Conservative  0;  Mismatches 304;  Indels  0;  Gaps  0;
```

```
Qy   230 CCGAGGAGGTGAACCGGCTCACGGAGGAGCAACCTACCGGAATGTTATGGAACAGTTCAATC 289
Db   112 CAGAGGAGATGCACCGGGCTCACGGAAAATGTCTATAAGACCATCATGGAGGACAGTTCAACC 171
Qy   290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATATTATGAGAAAGCTGTAAACGCTATGA 349
Db   172 CTAGCCTCCGGAACTTCATCGCCATGGGGGAAGAAATTACGAGAAGGCACTGGAGGTGTGA 231
Qy   350 TCCTGGCGAGAAAAGCCTACTACGATGGAGGTGGCCAAGATCGGTGAGATTGCCACTGGGGT 409
Db   232 CGTATGCAGCCAAAGGCTACTTTGACGCCCCTGGTGAAGATGGGGAGCTGGCCCAGGAGA 291
Qy   410 CCCCCGTGTCAACTGAACTGGGACATGTCCATAGAGATTTCAAGTACCCACAAGAAAC 469
Db   292 GCCAGGGCTCCAAAGAAACTCGGAGACGTTCTTCTTCCAGATGGCTAAGTCCACAGCCAGA 351
Qy   470 TCAACGAGAGTCTTGATGAAATATTTAAAAATTCCAAAGAGAATTATCCATGAGCTGG 529
Db   352 TCCAGAATCAGCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGG 411
Qy   530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAGAGATACCAAAACAG 589
Db   412 AGCAGAAGGTGGAGCTGGACTCCAGCTATCTGCCGCTAAAGAAAATACCAGACTG 471
Qy   590 AACACAAGAATAAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
Db   472 AGCAAAAGGACGAAGGCGACCCCTGGACAAGTGTCAGGCTGCTTACGCACGTCGGA 531
Qy   650 GGAAAAGCCAAGGAAGCCGGAAAGCGACACTCAAATATAGACACACAAGAAAATTGAGTATGTGG 709
Db   532 AGAAGAGCCAAGAGCCGAGAATCCTCAGAAGAGTACTCGGACAAGAGACTGCACTACATCG 591
Qy   710 AGACCGGTTACTTCTCGTCAGAGAGTGAAATCCAGAGAGTGAAATGGTTGCAAAGAGG 769
Db   592 ACGCCCATCAGCAACACAGCCAGGGCCAGCTTCTGGTTGGATAAGGCAGTTCAAGCG 651
Qy   770 CTCTGCTTGAAGAGAAGAGCGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGCTTTGCAA 829
Db   652 CACTGACAGAGGAGGTGCAGGCGCCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCGTGGCCA 711
Qy   830 ACCACCATACATTATTATCACTTACAGTCTGCGAGAACTACTGAATTCCAAGCTGCCTCGGT 889
Db   712 AGAACTCCGCGGGCCTACCCACTCCAAGGGAGACGCTGCCCGCAGAAGCTGCCCTGT 771
Qy   890 GGCAGGAGACCTGTGTTTGATGCCATCAAAAGTGCCAGAGAAATCATGAATATGATCGAAG 949
Db   772 GGGCAACAGCCCTGTCCCAGCGACCCGAAGATCCCCGGCGCGGCGCGTGCAGCTCATCGAGC 831
Qy   950 AAATAAAGACCCCAGCCTCTACCCCC 975
Db   832 AGGTGGCCACCAACGGCGCCACCCTC 857
```

```
Best Local Similarity   59.2%;  Pred. No. 1.6e-48;
Matches  442;  Conservative  0;  Mismatches 304;  Indels  0;  Gaps  0;
```

```
Qy   230 CCGAGGAGGTGAACCGGCTCACGGAGGAGCAACCTACCGGAATGTTATGGAACAGTTCAATC 289
Db   112 CAGAGGAGATGCACCGGGCTCACGGAAAATGTCTATAAGACCATCATGGAGGACAGTTCAACC 171
Qy   290 CTGGGCTGCGAAATTTAATAAACCTGGGGAAAAATATTATGAGAAAGCTGTAAACGCTATGA 349
Db   172 CTAGCCTCCGGAACTTCATCGCCATGGGGGAAGAAATTACGAGAAGGCACTGGAGGTGTGA 231
Qy   350 TCCTGGCGAGAAAAGCCTACTACGATGGAGGTGGCCAAGATCGGTGAGATTGCCACTGGGGT 409
Db   232 CGTATGCAGCCAAAGGCTACTTTGACGCCCCTGGTGAAGATGGGGAGCTGGCCCAGGAGA 291
Qy   410 CCCCCGTGTCAACTGAACTGGGACATGTCCATAGAGATTTCAAGTACCCACAAGAAAC 469
Db   292 GCCAGGGCTCCAAAGAAACTCGGAGACGTTCTTCTTCCAGATGGCTAAGTCCACAGCCAGA 351
Qy   470 TCAACGAGAGTCTTGATGAAATATTTAAAAATTCCAAAGAGAATTATCCATGAGCTGG 529
Db   352 TCCAGAATCAGCTGGAAGAAATGCTGAAGTCTTTTCACAACGAGCTGCTTACGCAGCTGG 411
Qy   530 AGAAGAAGATAGAACTTGACGTGAAATATATGAACGCAACTCTAAAAGAGATACCAAAACAG 589
Db   412 AGCAGAAGGTGGAGCTGGACTCCAGCTATCTGCCGCTAAAGAAAATACCAGACTG 471
Qy   590 AACACAAGAATAAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAA 649
Db   472 AGCAAAAGGACGAAGGCGACCCCTGGACAAGTGTCAGGCTGCTTACGCACGTCGGA 531
Qy   650 GGAAAAGCCAAGGAAGCCGGAAAGCGACACTCAAATATAGACACACAAGAAAATTGAGTATGTGG 709
Db   532 AGAAGAGCCAAGAGCCGAGAATCCTCAGAAGAGTACTCGGACAAGAGACTGCACTACATCG 591
Qy   710 AGACCGGTTACTTCTCGTCAGAGAGTGAAATCCAGAGAGTGAAATGGTTGCAAAGAGG 769
Db   592 ACGCCCATCAGCAACACAGCCAGGGCCAGCTTCTGGTTGGATAAGGCAGTTCAAGCG 651
Qy   770 CTCTGCTTGAAGAGAAGAGCGCGCTTCTGCTTTCTGGTTGATAAGCACTGTGCTTTGCAA 829
Db   652 CACTGACAGAGGAGGTGCAGGCGCCTTCTGCTTCCTGGTGGAGAAGCAGTGCGCGTGGCCA 711
```

```
RESULT 3
US-09-046-572-2
; Sequence 2, Application US/09046572
; Patent No. 6589935
; GENERAL INFORMATION:
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Baughn, Mariah
;  TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;  NUMBER OF SEQUENCES: 5
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Incyte Pharmaceuticals, Inc.
;  STREET: 3174 Porter Dr.
;  CITY: Palo Alto
;  STATE: CA
;  COUNTRY: USA
;  ZIP: 94304
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Diskette
;  COMPUTER: IBM Compatible
;  OPERATING SYSTEM: DOS
;  SOFTWARE: FastSEQ for Windows Version 2.0
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/046,572
;  FILING DATE: Filed Herewith
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:
;  FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Billings, Lucy J.
;  REGISTRATION NUMBER: 36,749
;  REFERENCE/DOCKET NUMBER: PF-0323-1 CIP
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 650-855-0555
;  TELEFAX: 650-845-4166
;  INFORMATION FOR SEQ ID NO: 2:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 2080 base pairs
;  TYPE: nucleic acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  IMMEDIATE SOURCE:
;  LIBRARY: BRSTNOT06
;  CLONE: 918158
US-09-046-572-2

  Query Match             12.4%;  Score 259.6;  DB 3;  Length 2080;
  Best Local Similarity   59.2%;  Pred. No. 1.6e-48;
  Matches  442;  Conservative  0;  Mismatches 304;  Indels  0;  Gaps  0;
```

```
Qy      230 CCGAGGAGGTGAACCGGCTTCACGGAGAGCGAAGTGACACCTACCGGAATCC 289
                |||| |||||||||||| ||||||||||||||||  ||||| |  ||||||||
Db      112 CAGAGGAGATGCACCGGCTCACGGAAAATGTCTATAAGACACCATCATGGAGGACAGTTCAACC 171

Qy      290 CTGGGCTGGGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACCGCTATGA 349
                |||||| ||||  ||||||||  |||  ||| || |||||||||||||||  |||| |||
Db      172 CTAGCCTCCGGAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCACTGGCAGGTGTGA 231
```

```
Qy      830 ACCACATACATTATTATCACTTACAGTCTCGAGAACTACTGAATTCCAAGCTGCCTCGGT 889
                ||||| || ||||||| |||||||||||||||||||||||  ||||||||||||||||
Db      712 AGAACTCCGGCGCCTACCACTCCAAGGCCAAGGAGTCGTGCTCCGCCAGCAGGAGCGGTCGT 771

Qy      890 GGCAGGAGGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAATCATGAATAGATCGAAG 949
                ||||  ||||| |||||||||||| |||||| |  ||||| |||||||| ||||||||
Db      772 GGCAACAGGCCCTGTCCGACCCCAGCAAGAGCGCGGGCGGGTGCAGCGTCATCGAGC 831

Qy      950 AAATAAAGACCCCAGCCCTCTACCCCC 975
                |||||||||||||||||||||||||||
Db      832 AGGTGGCCAGCAACGGCGCCACCCTC 857
```

```
RESULT 4
US-09-046-572-4
; Sequence 4, Application US/09046572
; Patent No. 6589935
; GENERAL INFORMATION:
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Baughn, Mariah
;  TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
;  NUMBER OF SEQUENCES: 5
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Incyte Pharmaceuticals, Inc.
;  STREET: 3174 Porter Dr.
;  CITY: Palo Alto
;  STATE: CA
;  COUNTRY: USA
;  ZIP: 94304
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Diskette
;  COMPUTER: IBM Compatible
;  OPERATING SYSTEM: DOS
;  SOFTWARE: FastSEQ for Windows Version 2.0
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/046,572
;  FILING DATE: Filed Herewith
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
```

```
Qy      350 TCCTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGAATCCGTGAGATTGCCACTCGGT 409
                 ||||                                                |||||||
Db      232 CGTATGCAGCCAAAAGGCTACTTTGACCCCGTGTGAAGATGGGAGCGCAGCAGGA 291

Qy      410 CCCCGTGTCAACTGACATGGGACATGTCCTCATAGAGAGATTTCAAGTACCACAGAAAC 469
                |||| ||||||||||||||||||||||| |||| ||||||||||  ||||||||||||
Db      292 GCCAGGGCTCCAAAGAACTCGGAGACGTTCTCTTCCACCAGATGGCTGAAGTCCACAGGCAGA 351

Qy      470 TCAACGAGAGTCTTGATGAAAAATTTAAAAAAATTCAACAAAGAGATTATCCATGAGCTGG 529
                ||||||||||||||||||   |  |||| ||  ||||||||||  |||| |||||||||
Db      352 TCCAGAATCAGCTGGAGAAAATGCTGAAAGTCTTTCACAACGAGGCTCTACGCGAGCTGG 411

Qy      530 AGAAGAAGATAGAACTTGACGTGAAATATATGAAGGCAAGCTCTAAAAAGATACCAAAC 589
                |||||||||||||||||||||| || |  |||||||| || |||||||||||||||||
Db      412 AGCAGAAGGTGGAGCTGGACTTACCGTATCTGAGTCGCTGGGTACCAGAGATCAGA 471

Qy      590 AACACAAGAGAATAAATTAGGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAGGAAGATCAGAAC 649
                ||||||||||||| ||  |||||||||||||||||||||||||||||||||||||||  |
Db      472 AGCAAAGGAGGAGGAAAAGGCCGAGGCCGCCCTGGACAGCTGGGAAGCGATGCTCGGA 531

Qy      650 GGAAAAGCCCAGGAAGCCGGAAACGGACTCAAATATGAACACAAAGAAAATTGAGTAGTGTGG 709
                ||| |||| ||||| || ||||  ||| ||||||||| |||||||||||||| | |||||
Db      532 AGAAGAGACCCAGGGCCAGCAAGAATCCTCAGAAGGTACTACGGACAAGGAGCCAGTACATCG 591

Qy      710 AGACCCGTTACTTCTCGTCAGAGTCGAGAATCCAGAAAATTCATTGCAGATCGGTTGCAAACAGG 769
                |||||| |  |||||| |||| |||||||| |||| |||||| | |||| || | |||||
Db      592 ACGCCCATCAGCAACACAGCCAGGCGAGCTGGGAGAATTACGTGTCCGACGGCTACCAAAGACG 651

Qy      770 CTCTGCTTGAAGAGAAGAGGCGCCTTCGTTCGTTTCGTGGTTGATTCGAGAAGACTGTGGCTTTGCAA 829
                |||||  ||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      652 CACTGACAGAGGAGGTGCAGGCGCCTTCGGCCTTGTCGTTTCCTGGTTGGAGAGCAGTGCGCGCCA 711

Qy      830 ACCACATACATTATTATCACTTACAGTCTCGAGAACTACTGAATTCCAAGCTGCCTCGGT 889
                ||||| || |||||||||||||||||||||||||||||||||  ||||||||||||||||
Db      712 AGAACTCCGGCGCCTACCACTCCAAGGCCAAGGAGTCGTGCTCCGCCAGCAGGAGCGGTCGT 771

Qy      890 GGCAGGAGGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAATCATGAATATGATCGAAG 949
                ||||  ||||| |||||||||||||| |||||| |  ||||| |||||| || |||||||||
Db      772 GGCAACAGGCCCTGTCCGACCCCAGCAAGAGCGCGGGCGGGTGCAGCGTCATCGAGC 831

Qy      950 AAATAAAGACCCCAGCCCTCTACCCCC 975
                |||||||||||||||||||||||||||
Db      832 AGGTGGCCAGCAACGGCGCCACCCTC 857
```

```
;  APPLICATION NUMBER: US/09/046,572
;  FILING DATE: Filed Herewith
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
```

```
;    APPLICATION NUMBER:
;    FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;    NAME:  Billings, Lucy J.
;    REGISTRATION NUMBER: 36,749
;    REFERENCE/DOCKET NUMBER:  PP-0323-1 CIP
;    TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  650-855-0555
;    TELEFAX:  650-845-4166
;    INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;    LENGTH:  2328 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  single
;    TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;    CLONE:  249150
;    LIBRARY:  ADRETUT05
US-09-046-572-4

Query Match      11.5%;    Score 241.6;  DB 3;  Length 2328;
Best Local Similarity    58.0%;  Pred. No. 1.9e-44;
Matches  427;  Conservative  0;  Mismatches 309;  Indels  0;  Gaps  0;

Qy        240 GAACCGGCCTCACGGAGGAGCACTACCGGATGTTAGGGACAGTTCAATCCTGGGGCTGCG 299
              |||||
Db        278 GAACTTTCGTGGTGAGAGTGGGCAGGGGACCCATCATGGAGCAGTTCAACCCTAGCCTCCG 337

Qy        300 AAAATTTAATAAACCTGGGGAGAAATTATGATGGAGAAAGCTGTAAACGCTATAGATCCTGGCCAGG 359
              |||||  ||||  ||||
Db        338 GAACTTCATCGCCATGGGGAAGAATTACGAGAAGGCCACTGGCAGGTGTGAGGTATCCAGC 397

Qy        360 AAAAGCCTACTACGATGGAGTGGGCCAAGATCGGTGAGATTGCCACTGGGTCTCCCCGTGTC 419
              ||  ||||   |||
Db        398 CAAAGGGCTACTTTGACGCCCTGGTCGAAGATGGGGGAGCTGGCCCACGGAGCCAGGGCTC 457

Qy        420 AACTGAACTGGGACAATGTCCTCCATAGAGATTTCAAGTACTACCACCAAGAAACTCAACGAGAG 479
              |||||   ||||  ||| ||||
Db        458 CAAAGAACTCGGAGACGTTCTCTTCCCAGATGGCTGAAGTCCACAGGCCAGATCCAGAATCA 517

Qy        480 TCTTGATGAAAATTTTAAAAAATTCCACAAAAGAGATTATCCATGAGCTGGAGAAGAGAT 539
              ||||||  ||||  ||| |||
Db        518 GCTCGAAGAAAATGCTGAAGTCTTTTCACAACGAGCTGTGCTTACGCAGCTGAGCGAGAAGGT 577

Qy        540 AGAACTTGACGTGAAATATATGAACGCAACTCTAAAAAGATACCAAAACAGAACACAAGA 599
              ||| |||
Db        578 GGAGCTGGACTCCAGGTATCTGAGTGCTCCGCTCGAAGAAAATACCGAGACTGAGCAAAGGAG 637

Qy        600 TAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAGAAGGAAAAGCCA 659
              |||| |||  |||||
Db        638 CAAAGGCGACGCCCCTGGACAAAGTGTCAGGCTGAGCGAAGAGCTTCGGAGAAGAGAGCCA 697

Qy        660 AGGAAGCCGAAACCCACTCAAATATGAACACAAGAAAATTGAGATTATGTGAGACCGTTAC 719
              ||| ||||   |||
Db        698 GGGCAGCAAGAATCTTCAGAAGTACTCGGACAAGAGAGCTCGAGTCAATCGAGCCCATCAG 757

Qy        720 TTCTCGTCAGAGTGAAATCCAGAAAATTCATTGCAGAATGGTTGCAAAGAGGCTTCGCTTGA 779
              |||  |||
Db        758 CAACAAGCCAGGGCGAGCTGGAGAATTACGTGTCCGACAGCACCGCACTGACACAGA 817

Qy        780 AGAGAAGAGAGCGCGCCTTCTGCTTTGTGATAAGCACTGTCGGCTTTGCAAACCACCATACA 839
              |||  |||  ||||  ||||  |||
Db        818 GGAGCGCGCAGCGCCGCTTCCTGCGCTTCCCGTGGAGAAGCAGTGCCCCGTGGCCAAGAACTCCGC 877

Qy        840 TTATTATCACTTACAGTCTGCGCAGAACTACTGAATCCCAAGCTGCTCGGTGGCCAGAGAC 899
              |||  |||
Db        878 GGCCTACCACTCCAAGGGCAAGGAGCTGCTCCCGCAGAAGAAATCATGAATATGATCGAAGAGAAATAAAGAC 937

Qy        900 CTGTGTTGATGCCATCAAAAGTGCCAGAAAAATCATGAATATGATCGAAGAGAAATAAAGAC 959
              ||||
Db        938 CTGTGCCGACACCCACGCAAGATCCCGGAGCGCGGCGGCTGCAGCCAGCAGGTGCCCAG 997

Qy        960 CCCAGGCTCTACCCCC 975
              |||||
```

;   RESULT 5
;   US-09-907-794A-376
;   Sequence 376, Application US/09907794A
;   Patent No. 6635468
;   GENERAL INFORMATION:
;   APPLICANT: Genentech, Inc.
;   APPLICANT: Ashkenazi, Avi
;   APPLICANT: Botstein, David
;   APPLICANT: Desnoyers, Luc
;   APPLICANT: Eaton, Dan L.
;   APPLICANT: Ferrara, Napoleone
;   APPLICANT: Filvaroff, Ellen
;   APPLICANT: Fong, Sherman
;   APPLICANT: Gao, Wei-Qiang
;   APPLICANT: Gerber, Hanspeter
;   APPLICANT: Gerritsen, Mary E.
;   APPLICANT: Goddard, A.
;   APPLICANT: Godowski, Paul J.
;   APPLICANT: Grimaldi, Christopher J.
;   APPLICANT: Gurney, Austin L.
;   APPLICANT: Hillan, Kenneth J.
;   APPLICANT: Kljavin, Ivar J.
;   APPLICANT: Mather, Jennie P.
;   APPLICANT: Pan, James
;   APPLICANT: Paoni, Nicholas F.
;   APPLICANT: Roy, Margaret Ann
;   APPLICANT: Stewart, Timothy A.
;   APPLICANT: Tumas, Daniel
;   APPLICANT: Williams, P. Mickey
;   APPLICANT: Wood, William I.
;   TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;   TITLE OF INVENTION: Acids Encoding the Same
;   FILE REFERENCE: 10466-14
;   CURRENT APPLICATION NUMBER: US/09/907,794A
;   CURRENT FILING DATE: 2001-07-17
;   PRIOR APPLICATION NUMBER: PCT/US00/04414
;   PRIOR FILING DATE: 2000-02-22
;   PRIOR APPLICATION NUMBER: US 60/143,048
;   PRIOR FILING DATE: 2000-01-05
;   PRIOR APPLICATION NUMBER: US 60/145,698
;   PRIOR FILING DATE: 1999-07-07
;   PRIOR APPLICATION NUMBER: US 60/146,222
;   PRIOR FILING DATE: 1999-07-26
;   PRIOR APPLICATION NUMBER: US 60/146,222
;   PRIOR FILING DATE: 1999-07-28
;   PRIOR APPLICATION NUMBER: PCT/US99/20594
;   PRIOR FILING DATE: 1999-09-08
;   PRIOR APPLICATION NUMBER: PCT/US99/20944
;   PRIOR FILING DATE: 1999-09-13
;   PRIOR APPLICATION NUMBER: PCT/US99/21090
;   PRIOR FILING DATE: 1999-09-15
;   PRIOR APPLICATION NUMBER: PCT/US99/21547
;   PRIOR FILING DATE: 1999-09-15
;   PRIOR APPLICATION NUMBER: PCT/US99/23089
;   PRIOR FILING DATE: 1999-10-05
;   PRIOR APPLICATION NUMBER: PCT/US99/28214
;   PRIOR FILING DATE: 1999-11-29
;   PRIOR APPLICATION NUMBER: PCT/US99/28313
;   PRIOR FILING DATE: 1999-11-30
;   PRIOR APPLICATION NUMBER: PCT/US99/28564
;   PRIOR FILING DATE: 1999-12-02
;   PRIOR APPLICATION NUMBER: PCT/US99/28565
;   PRIOR FILING DATE: 1999-12-02
;   PRIOR APPLICATION NUMBER: PCT/US99/30095
;   PRIOR FILING DATE: 1999-12-16
;   PRIOR APPLICATION NUMBER: PCT/US99/30911
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: PCT/US99/30999
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: PCT/US00/00219
;   PRIOR FILING DATE: 2000-01-05

```
Db  998 CAACGGGCGCCACCCTC 1013
        |  |   ||  |||||  |
```

; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-376

Query Match
Best Local Similarity 81.4%; Score 82.8; DB 3; Length 997;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Pred. No. 7.4e-09;

```
Qy   1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTGGGTTTCAATCTTTAAAAAAAA 2038
          |||.||||||||.||||||||||||| |||||||||||.||||| |||||||||||||
Db    869 ATAAGTATTTGTTACTCTGATACAAAATAAAAATAAGTAGTTTTTAAATGTTAAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  986
```

RESULT 6
US-09-905-125A-376
; Sequence 376, Application US/0905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-376

Query Match
Best Local Similarity 81.4%; Score 82.8; DB 3; Length 997;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Pred. No. 7.4e-09;

```
Qy   1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTGGGTTTCAATCTTTAAAAAAAA 2038
          |||.||||||||.||||||||||||| |||||||||||.||||| |||||||||||||
Db    869 ATAAGTATTTGTTACTCTGATACAAAATAAAAATAAGTAGTTTTTAAATGTTAAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  986
```

RESULT 7
US-09-902-775A-376
; Sequence 376, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22

APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,700
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-906-700-376

Query Match   4.0%;   Score 82.8;   DB 3;   Length 997;
Best Local Similarity   81.4%; Pred. No. 7.4e-09;
Matches   96; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   1979 ATATTTTCTTCCCTATTCTCGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAT 2038
                  ||| |||||||| ||||| ||||||||||||||||| ||||| |||||| |||||||||
Db   869 ATAAGTATTTGTTACTCTGATACACAAATAAAAATAAGTAGTTTTTAAAATGTTAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 9
US-09-903-603A-376
Sequence 376, Application US/09903603A
Patent No. 6767995
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.

RESULT 8
US-09-906-700-376
Sequence 376, Application US/09906700
Patent No. 6723535
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.

Query Match   4.0%;   Score 82.8;   DB 3;   Length 997;
Best Local Similarity   81.4%; Pred. No. 7.4e-09;
Matches   96; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   1979 ATATTTTCTTCCCTATTCTCGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAT 2038
                  ||| |||||||| ||||| ||||||||||||||||| ||||| |||||| |||||||||
Db   869 ATAAGTATTTGTTACTCTGATACACAAATAAAAATAAGTAGTTTTTAAAATGTTAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

Query Match   4.0%;   Score 82.8;   DB 3;   Pred. No. 7.4e-09;   Length 997;
Best Local Similarity   81.4%;

US-09-903-603A-376
ORGANISM: Homo Sapien
TYPE: DNA
LENGTH: 997
SEQ ID NO 376
NUMBER OF SEQ ID NOS: 423

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: GNE.1618P2C12
; CURRENT APPLICATION NUMBER: US/09/903,603A
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05

Matches   96;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

```
Qy   1979  ATATTTCTTCCTATTCTCTGTTCAAGAAACAGTAAACTGGTTTCAATCTTTAAAAAAAA  2038
           |||||  ||| ||| ||| ||||||||||||  |  |||||||  || |||||||||||
Db    869  ATAAGTATTTGTTACTCTGATACAAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAA   928

Qy   2039  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2096
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    929  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA        986
```

RESULT 10
US-09-904-920A-376
; Sequence 376, Application US/0990420A
; Patent No. 6806352
GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,920A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16

RESULT 11

```
US-09-909-064-376
Sequence 376, Application US/09909064
Patent No. 6818449
GENERAL INFORMATION:
```

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/909,064
CURRENT FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944

```
US-09-904-920A-376
SEQ ID NO 376
  LENGTH: 997
  TYPE: DNA
  ORGANISM: Homo Sapien
NUMBER OF SEQ ID NOS: 423
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-16
```

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy  1979  ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA  2038
          || |   || |||| ||||||||||||||||||||||||||| |||| ||||||||||||
Db   869  ATAAGTATTTGTTACTCTGATACACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAA  928

Qy  2039  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  986
```

RESULT 12

```
US-09-905-381A-376
Sequence 376, Application US/09905381A
Patent No. 6818746
GENERAL INFORMATION:
```

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.

```
US-09-909-064-376
SEQ ID NO 376
  LENGTH: 997
  TYPE: DNA
  ORGANISM: Homo Sapien
NUMBER OF SEQ ID NOS: 423
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-13
```

Query Match 4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy  1979  ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA  2038
          || |   || |||| ||||||||||||||||||||||||||| |||| ||||||||||||
Db   869  ATAAGTATTTGTTACTCTGATACACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAA  928

Qy  2039  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  986
```

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,381A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/145,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien

US-09-905-381A-376

Query Match   4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
  Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1979 ATATTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
              ||| |||||||||||||||| | |||||||| ||||||||||||||||||||||||||||
Db   869 ATAAGTATTTGTTACTCTGATACAAATAAAATAAGTAGTTTTTAAAATGTTAAAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 13
US-09-906-618-376
; Sequence 376, Application US/0990618
; Patent No. 6828146
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,618
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/145,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien

US-09-906-618-376

Query Match   4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
  Matches 96; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1979 ATATTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
              ||| |||||||||||||||| | |||||||| ||||||||||||||||||||||||||||
Db   869 ATAAGTATTTGTTACTCTGATACAAATAAAATAAGTAGTTTTTAAAATGTTAAAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 14
US-09-906-646-376
; Sequence 376, Application US/09906646
; Patent No. 6852848
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,646
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376

; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-906-646-376

Query Match   4.0%; Score 82.8; DB 3; Length 997;
Best Local Similarity 81.4%; Pred. No. 7.4e-09;
Matches   96; Conservative   0; Mismatches   22; Indels   0; Gaps   0;

Qy   1979 ATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAAA 2038
          ||| ||||||||| || | ||||||||||||||| |||| | ||||||| || ||||||||
Db   869 ATAACTATTTTGTTACTTCTGATACAAAATAAGTAGTTTTAAAATGTTAAAAAAAAA 928

Qy   2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 986

RESULT 15
US-09-904-462-376
; Sequence 376, Application US/09904462
; Patent No. 6878807
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,462
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-1B
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 2000-09-1B
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

```
;   PRIOR APPLICATION NUMBER: PCT/US99/28214
;   PRIOR FILING DATE: 1999-11-29
;   PRIOR APPLICATION NUMBER: PCT/US99/28313
;   PRIOR FILING DATE: 1999-11-30
;   PRIOR APPLICATION NUMBER: PCT/US99/28564
;   PRIOR FILING DATE: 1999-12-02
;   PRIOR APPLICATION NUMBER: PCT/US99/28565
;   PRIOR FILING DATE: 1999-12-02
;   PRIOR APPLICATION NUMBER: PCT/US99/30095
;   PRIOR FILING DATE: 1999-12-16
;   PRIOR APPLICATION NUMBER: PCT/US99/30911
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: PCT/US99/30999
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: PCT/US99/30999
;   PRIOR FILING DATE: 1999-12-20
;   PRIOR APPLICATION NUMBER: PCT/US00/00219
;   PRIOR FILING DATE: 2000-01-05
;   NUMBER OF SEQ ID NOS: 423
;   SEQ ID NO 376
;     LENGTH: 997
;     TYPE: DNA
;     ORGANISM: Homo Sapien
US-09-904-462-376

Query Match             4.0%;  Score 82.8;  DB 3;  Length 997;
Best Local Similarity  81.4%;  Pred. No. 7.4e-09;
Matches  96;  Conservative  0;  Mismatches  22;  Indels  0;  Gaps  0;

Qy    1979 ATATTTTCTTCCTATTTCTGTTCAAGAAACAGTAAACTTGGTTTCAATCTTTAAAAAAAA 2038
              || |||  || ||| |||||:|||  |||  ||||| | |||| ||||||||||||
Db     869 ATAAGTATTTGTTACTCTGATACACAATAAAAATAAGTAGTTTTTAAATGTTAAAAAAAA  928

Qy    2039 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2096
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  986

Search completed: March 20, 2006, 16:06:09
Job time : 382 secs
```

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.7
          Copyright (c) 1993 - 2006  Bioacceleration Ltd.
```

Run on:　　　　　March 20, 2006, 13:40:03 ; Search time 1705 seconds
　　　　　　　　　　　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　　　　　　　　　　10165.757 Million cell updates/sec

Title:　　　　　2096
Perfect score:　2096
Sequence:　　　1 ctccctcgctcctcgaagaa............aaaaaaaaaaaaaaaaaaaa 2096

Scoring table:　IDENTITY_NUC
　　　　　　　　　Gapop 10.0 , Gapext 1.0

Searched:　　　9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters:　　19587084

Minimum Match length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
　　　　　　　　　Maximum Match 100%
　　　　　　　　　Listing first 45 summaries

Database :　　Published Applications NA Main:*
```
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 6: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 8: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description
1	2096	100.0	2096	9	US-10-801-292-1	Sequence 1, Appli
2	2025.2	96.6	2527	9	US-10-756-149-4599	Sequence 4599, Ap
3	1514.6	72.3	2471	7	US-10-112-944-39	Sequence 39, Appl
4	1448.4	69.1	1510	7	US-10-112-944-523	Sequence 523, App
5	1216.4	58.0	2120	6	US-10-172-118-1805	Sequence 1805, Ap
6	1216.4	58.0	2120	7	US-10-342-887-1805	Sequence 1805, Ap
7	1071	51.1	1729	9	US-10-450-763-9761	Sequence 9761, Ap
8	1021.2	48.7	1993	9	US-10-450-763-9762	Sequence 9762, Ap
9	594	28.3	625	5	US-10-066-543-481	Sequence 481, App
10	576.4	27.5	1242	5	US-09-925-301-269	Sequence 269, App
11	576.2	27.5	641	5	US-10-066-543-990	Sequence 990, App
12	556	26.5	559	3	US-09-878-134-252	Sequence 252, App
13	520.8	24.8	559	5	US-10-066-543-321	Sequence 321, App
14	488	23.3	511	3	US-09-960-253-83	Sequence 83, Appl
15	470.4	22.4	512	9	US-09-960-253-90	Sequence 90, Appl
16	329	15.7	449	9	US-10-450-763-9759	Sequence 9759, Ap
17	274.8	13.1	439	9	US-10-450-763-9760	Sequence 9760, Ap
18	259.6	12.4	2080	6	US-10-435-324-2	Sequence 2, Appli
19	255.5	12.2	295	6	US-09-867-701-2669	Sequence 2669, Ap
20	241.6	11.5	2328	6	US-10-435-324-4	Sequence 4, Appli
21	240.2	11.5	45698	3	US-09-984-429-344	Sequence 344, App
22	227.4	10.8	1415	6	US-10-970-760-1	Sequence 1, Appli
23	216.4	10.3	16181	3	US-09-764-847-1426	Sequence 1426, Ap
24	216.4	10.3	16181	3	US-09-764-891-6956	Sequence 6956, Ap
25	216.4	10.3	16181	5	US-10-092-154-1426	Sequence 1426, Ap
26	209.6	10.0	476	3	US-09-964-824A-78	Sequence 78, Appl
27	209.6	10.0	476	3	US-09-964-824A-491	Sequence 491, Appl
28	209.6	10.0	476	3	US-09-967-768A-205	Sequence 205, App
29	209.6	10.0	476	9	US-10-843-641A-5381	Sequence 5381, Ap
30	209.6	10.0	476	9	US-10-843-641A-5794	Sequence 5794, Ap.
31	209.6	10.0	476	9	US-10-843-641A-6350	Sequence 6350, Ap
32	203.8	9.7	363	9	US-10-313-669-23	Sequence 23, Appl
33	199	9.5	277	9	US-10-779-543-6174	Sequence 6174, Ap
34	187	8.9	454	7	US-10-764-891-820	Sequence 820, App
35	184	8.8	490	6	US-09-876-143-1286	Sequence 1286, Ap
36	176	8.4	176	3	US-09-867-701-2765	Sequence 2765, Ap
37	158.4	7.6	2316	6	US-10-094-749-1190	Sequence 1190, Ap
38	157	7.5	996	3	US-10-027-632-101199	Sequence 101199,
39	152	7.3	2242	5	US-10-027-632-102200	Sequence 102200,
40	152	7.3	2242	5	US-10-027-632-102200	Sequence 102200,
41	152	7.3	2242	5	US-10-027-632-102201	Sequence 102201,
42	152	7.3	2242	6	US-10-027-632-102199	Sequence 102199,
43	152	7.3	2242	6	US-10-027-632-102200	Sequence 102200,
44	152	7.3	2242	6	US-10-027-632-102201	Sequence 102201,
45	147.2	7.0	640	4	US-09-925-065A-730687	Sequence 730687,

ALIGNMENTS

```
RESULT 1
US-10-801-292-1
; Sequence 1, Application US/10801292
; Publication No. US20050202448A1
; GENERAL INFORMATION:
; APPLICANT: LEE, YI-CHAO
; APPLICANT: YUEN, PUI-YEE
; APPLICANT: HUANG, YI-HUEI
; APPLICANT: WU, HUI-CHUAN
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF CANCER
; FILE REFERENCE: 5422-2
; CURRENT APPLICATION NUMBER: US/10/801,292
; CURRENT FILING DATE:  2004-03-15
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 3.2
;
; SEQ ID NO 1
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-801-292-1
```

Query Match　　　　　　100.0%;　Score 2096;　DB 9;　Length 2096;
Best Local Similarity　100.0%;
Matches 2096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTCCCTCGCTCCTCGAAGAAGAGGCCAGGGCGGGGGGCTGCCCGAGACATTTTGGCAG   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCCTCGCTCCTCGAAGAAGAGGCCAGGGCGGGGGGCTGCCCGAGACATTTTGGCAG   60

Qy         61 CGGAGACGCGCGGCGGGCCGACACTCTCGGGGCCGACGGGCTTCGGCGCGGCCGACCCTCCAGAGC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGAGACGCGCGGCGGGCCGACACTCTCGGGGCCGACGGGCTTCGGCGCGGCCGACCCTCCAGAGC  120

Qy        121 CCCTTAGTTCGGCGGCCCCGGGCCCCTCCCCGCTGCCCGGGGAGTCCGGGGCGCGGGCCACGGCCG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTTAGTTCGGCGGCCCCGGGCCCCTCCCCGCTGCCCGGGGAGTCCGGGGCGCGGGCCACGGCCG  180

Qy        181 CGTCCTCCCCGGCGCTTGCTTCCGCCCGGGCCCGGGGGCCGCAGCCAATGTCCCGGGGGGGTG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTCCTCCCCGGCGCTTGCTTCCGCCCGGGCCCGGGGGCCGCAGCCAATGTCCCGGGGGGGTG  240

Qy        241 AACCGGCTCACGGAGGAGCACCTACCGGAAATGTTATGGAACAGTTCAATCCTGGGGTCGCGA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCGGCTCACGGAGGAGCACCTACCGGAAATGTTATGGAACAGTTCAATCCTGGGGTCGCGA  300
```

```
Qy    301  AATTTAATTAAAAACCTGGGGGAAAAAATTATGTGAGAAAGCTGTAAACGCTATGTCCTGGCAGGA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AATTTAATTAAAAACCTGGGGGAAAAAATTATGTGAGAAAGCTGTAAACGCTATGTCCTGGCAGGA  360

Qy    361  AAAGCCTACTACGATGGAGTGGGCCAAGATGGGTGAGATTGGCCACTGGGTGTCCCCGTGTCA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AAAGCCTACTACGATGGAGTGGGCCAAGATGGGTGAGATTGGCCACTGGGTCCCCCTGTCA  420

Qy    421  ACTGAAACTGGGACACATGTTCTTCATAGAGATTTCAGTCACACAGAAACTTCAACAGAGT  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GACTTGACGTGAAATATATGAAATTCAAGTACCCACCACAGAAGTGTTCCATGGGGGCAGCT  480

Qy    481  CTTGATGAAAATTTTAAAAATTCCACAAAGAGATTATCCATGGATGGAGGAGAGATA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CTTGATGAAAATTTTAAAAATTCCACAAAGAGATTATCCATGGATGGAGGAGAGATA  540

Qy    541  GAACTTGACGTGAAATATATGAAACGGAACTTCTAAAAGATACCAACAGAACCAAGAAT  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GAACTTGACGTGAAATATATGAAACGGAACTTCTAAAAGATACCAACAGAACCAAGAAT  600

Qy    601  AAATTAGATCTTTGGAGAAATATGGAGAAATCCCAAGCTCAAAAAGAAGAAAGCCAA  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  AAATTAGATCTTTGGAGAAATCCCAAGCTCAAAAAGAAGAAAGCCAA  660

Qy    661  GGAAGCCGAAACGGCACTCAAAATATGAACAAAAGAAAAATTGAGTATGTGGGAGGAGCTCTTACT  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GGAAGCCGAAACGGCACTCAAAATATGAACAAAAGAAAAATTGAGTATGTGGGAGGAGCTCTTACT  720

Qy    721  TCTCGTCAGAGTGAAAATCCAGAAATTCATTGCAGAATGGTTGCAAAAGAGGCTTCTGCTTGAA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GAGAAGGCGGGCTTCTGCTTTCTCGGTTGATAAGCACTGTGCAAACCACATACAT  780

Qy    781  GAGAAGGCGGGCTTCTGCTTTCTCGGTTGATAAGCACTGTGCAAACCACATACAT  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  TATTATCACTTACAGTCTCGAGAACTACCAAGCTGCCTCGGTGCCAGGAGACC  840

Qy    841  TATTATCACTTACAGTCTCGAGAACTACCAAGCTGCCTCGGTGCCAGGAGACC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATGGAGAAAATAAAGACC  900

Qy    901  TGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATGGAGAAAATAAAGACC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CCAGCCTCTACCCCGTGTCTGGAACTCCTCACGGCTTCACCCATGATCGAGAGAGCAAT  960

Qy    961  CCAGCCTCTACCCCGTGTCTGGAACTCCTCACGGCTTCACCCATGATCGAGAGAGCAAT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GTGGGTTAGGAGAAAGATTACGACACACCCTTTCTAAAATGCTCACCAAAGATGCCCCCTCCT  1020

Qy   1021  GTGGGTTAGGAGAAAGATTACGACACACCCTTTCTAAAATGCTCACCAAAGATGCCCCCTCCT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  TCAGGCCAGAGCATATACCAGTCCCTTGATCGTAATGAATGCCAGCCACGGCTGCC  1080

Qy   1081  TCAGGCAGAGCATATACCAGTCCCTTGATCGTAATGAATGCCAGCCACGGCTGCC  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CCGAATTCACAAAGGGTAAAATAATTCAACAGGTACTTCCGAAAGATCCCAGTTACAGCGA  1200

Qy   1141  CCGAATTCACAAAGGGTAAAATAATTCAACAGGTACTTCCGAAAGATCCCAGTTACAGCGA  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  TCAGGCCAGAGCATATACCAGTCCCTTGATCGTAATGAATGTTTAATAACCCAGCCACGGCTGCC  1200

Qy   1201  TCAGTTTCGGTTGCAAACGGGACTGAACATGATGAAGAAGCAGAGAAGTGAAGACCA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TCAGTTTCGGTTGCAAACGGGACTGAACATGATGAAGAAGCAGAGAAGTGAAGACCA  1260

Qy   1261  CCGCACACTGCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAGTCATCACG  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  CCGCACACTGCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAGTCATCACG  1320

Qy   1321  CTGCTCATCCCCGAGGGAGGAGAGGAATGGCCTGGCTCTCTATGGAGAGAAGTCCAAGGCG  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  CTGCTCATCCCCGAGGGAGGAGAGGAATGGCCTGGCTCTCTATGGAGAGAAGTCCAAGGCG  1380

Qy   1381  AGGGGTTGGTTCCCGTCGTGGTACACGAGTTGCTGGAAGAAAATGAGACGAAGCAGTG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  AGGGGTTGGTTCCCGTCGTGGTACACGAGTTGCTGGAAGAAAATGAGACGAAGCAGTG  1440

Db   1381  ACCGGTGCCCACGCCAAGGCCCCACACCACAGCCACCGTGAAAGTGAGACTTGTCTGAG  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441  ACCGGTGCCCACGCCAAGGCCCCACACCACAGCCACCGTGAAAGTGAGACTTGTCTGAG  1500

Db   1441  AATAGCAGTGTTGTCATCCCCACACCGACTACTTGGAATGCTTGTCCATGGGGGGGCAGCT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501  AATAGCAGTGTTGTCATCCCCACACCGACTACTTGGAATGCTTGTCCATGGGGGGGCAGCT  1560

Db   1501  GCCGACGAGGAGGACAGATTCGGCACGAGACGACATCCACCTTTAAGGCCCCAGTCCAAG  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561  GCCGACGAGGAGGACAGATTCGGCACGAGACGACATCCACCTTTAAGGCCCCAGTCCAAG  1620

Db   1561  CCCGAGACCCGGGCTCCTAACGATGCCAAGGGAGACTGCAAAGCCGCCTTTTCTCAGCGGA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621  CCCGAGACCCGGGCTCCTAACGATGCCAAGGGAGACTGCAAAGCCGCCTTTTCTCAGCGGA  1680

Db   1681  GAAAACCCTTTGCCACTGTGAAAACTCCGCCGGACTGTGAAAAATGAATGCGGAGTCCCCT  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681  GAAAACCCTTTGCCACTGTGAAAACTCCGCCGGACTGTGAAAAATGAATGCGGAGTCCCCT  1740

Db   1741  ATCAATCGATGAGAGGACAGCACAGCCAAGGAGAACTCTCCCGCCGACTGTGACCAAGCAT  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741  ATCAATCGATGAGAGGACAGCACAGCCAAGGAGAACTCTCCCGCCGACTGTGACCAAGCAT  1800

Db   1801  ATGATGGGCGCCATCCTGTTCTGCCACGGTGCCACCAGCTTCAGGTAGTTGGGGTTGGGGAC  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1801  ATGATGGGCGCCATCCTGTTCTGCCACGGTGCCACCAGCTTCAGGTAGTTGGGGTTGAC  1860

Db   1861  AACTCTAAGTGTCGCGCCTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGGTTGAC  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1861  AACTCTAAGTGTCGCGCCTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGGTTGAC  1920

Db   1921  AAGTTTCTGCCTTTAAGATAAAATGAGTAAAAGTCTAAATGACCAGCCATTTAAAAT  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921  AAGTTTCTGCCTTTAAGATAAAATGAGTAAATGACCAGCCATTTAAAAT  1980

Db   1981  ATTTTCTTCCTATTCTGTTCAAGAAAACAGTAAACTTGGTTTCAAATCTTTAAAAAAAAT  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1981  ATTTTCTTCCTATTCTGTTCAAGAAAACAGTAAACTTGGTTTCAAATCTTTAAAAAAAAT  2040

Qy   2041  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2096
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2096
```

Best Local Similarity 99.9%; Conservative 0; Pred No. 0; Mismatches 3; Indels 0; Gaps 0;
Matches 2027;

```
Qy      1 CTCCTCTGCTTCCTCGAAGAAGGCAAGTTTTGACATTTTCGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCCTCTGCTTCCTCGAAGAAGGCAAGTTTTGACATTTTCGCAG 60

Qy     61 CGGAGACGCGGCGGGGGCACTCTCGGGCGCGGGGCGGGGCGGGGCTG 120
          |||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGGAGACGCGGCGGGGGCACTCTCGGGCGCGGGGCGGGGCGGGGCTG 120

Qy    121 CCCTTAGTCGTGGCGCCCCGCCCTCCCGCTGCCGCCCAGGCCG 180
          |||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCTTAGTCGTGGCGCCCCGCCCTCCCGCTGCCGCCCAGGCCG 180

Qy    181 CGTCCTCCCGGCGTTGCTCGGCCCCGGGGGCCCACCATGTCCCG 240
          ||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGTCCTCCCGGCGTTGCTCGGCCCCGGGGGCCCACCATGTCCCG 240

Qy    241 AACCGGCTCACGGAGCAGCCTACCGGAATGTTATCGGAACAGTTCAATCCTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AACCGGCTCACGGAGCAGCCTACCGGAATGTTATCGGAACAGTTCAATCCTGG 300

Qy    301 AATTTAATTAAAACCTGGGAGGAAAAATTATGAGAAACCTGTAAAACGGTA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AATTTAATTAAAACCTGGGAGGAAAAATTATGAGAAACCTGTAAAACGGTA 360

Qy    361 AAAGCCTACTACCGATGGGACTGGCCAAGATCGGGTGAGATTGGCCAAGATC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGCCTACTACCGATGGGACTGGCCAAGATCGGGTGAGATTGGCCAAGATC 420

Qy    421 ACTGAACTGGGACATGTCCTCATAGAGATTCAAGTACCCAAGAAACTCAACGAGAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACTGAACTGGGACATGTCCTCATAGAGATTCAAGTACCCAAGAAACTCAACGAGAGT 480

Qy    481 CTTGATGAAAAATTTTCCACAAGAGGATTATCCATGACTGGAGATGGAGGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGATGAAAAATTTTCCACAAGAGGATTATCCATGACTGGAGATGGAGGA 540

Qy    541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAGATACCAAAACGAACAAGAAT 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAACTTGACGTGAAATATATGAACGCAACTCTAAAAGATACCAAAACGAACAAGAAT 600

Qy    601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAGATCAGAAGGAAAGCCAA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAATTAGAGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAGATCAGAAGGAAAGCCAA 660

Qy    661 GGAAGCCGAAAACGCACTCAAATATGAACAAAAGAAATTGAGTAGTGTGGAGACCGTTACT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGAAGCCGAAAACGCACTCAAATATGAACAAAAGAAATTGAGTAGTGTGGAGACCGTTACT 720

Qy    721 TCTCGTCAGAGTGAAATTCAGTTGTTGCAAAGGAGGCTCTGCTTGAA 780
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCTCGTCAGAGTGAAATTCAGTTGTTGCAAAGGAGGCTCTGCTTGAA 780

Qy    781 GAGAGAGAGGGGCTTCTGCGTTCGGTTGATAAGGCACTGTGGTTGGTTGGCTTTGAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAGAGAGGGGCTTCTGCGTTCGGTTGATAAGGCACTGTGGTTGGTTGGCTTTGAC 840

Qy    841 TATTATCACTTACAGTGCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGGCAGGACC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TATTATCACTTACAGTGCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGGGCAGGACC 900

Qy    901 TGTGTTGATGCCATCAAAAGTGCTCTGAACTCCTCAGGCTTCACCCATGATCGAGAAGCAAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TGTGTTGATGCCATCAAAAGTGCTCTGAACTCCTCAGGCTTCACCCATGATCGAGAAGCAAT 960

Qy    961 CCAGCCTTCTACCCCCGTGTCTCAGGCTTCACGGCTT 1020
          |||||||||||||||||||||||||||||||||||||
Db    961 CCAGCCTTCTACCCCCGTGTCTCAGGCTTCACGGCTT 1020

Qy   1021 GTGGTTAGGGAAAGATTACGACACACCCTTTCTAAATGCTACCAAAGATGCCCCCCGCTCCT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTGGTTAGGGAAAGATTACGACACACCCTTTCTAAATGCTACCAAAGATGCCCCCCGCTCCT 1080
```

```
Db   1021 GTGGTTAGGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCT 1080
Qy   1081 TCAGGCAGGACCATATACCAGTCCCTTGATCGATATGTTTAATAACCAGCCACGGCTGCC 1140
Db   1081 TCAGGCAGGACCATATACCAGTCCCTTGATCGATATGTTTAATAACCAGCCACGGCTGCC 1140

Qy   1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTACGCGGA 1200
Db   1141 CCGAATTCACAAAGGGTAAATAATTCAACAGGTACTTCCGAAGATCCCAGTTACGCGGA 1200

Qy   1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260
Db   1201 TCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTC 1260

Qy   1261 CCGCACACTCCGGGGCTTCCAACAAAGACCTTACTCACTTTCCACAGGGAGGATGTCATCACG 1320
Db   1261 CCGCACACTCCGGGGCTTCCAACAAAGACCTTACTCACTTTCCACAGGGAGGATGTCATCACG 1320

Qy   1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTCATGGGAGTCGTTGTTCCAAGGCG 1380
Db   1321 CTGCTCATCCCCGAGGAGAAGGATGGCTGGCTCTCATGGGAGTCGTTGTTCCAAGGCG 1380

Qy   1381 AGGGGTTGGTTCCCGTGCTGCTACAAGAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440
Db   1381 AGGGGTTGGTTCCCGTGCTGCTACAAGAGTTGCTGGAAGAAAATGAGACAGAAGCAGTG 1440

Qy   1441 ACCCTTGCCCACGCCAAGCCCCACCACCAGTGAAGCCATCACGACCGTGAACTTGTCTGAG 1500
Db   1441 ACCCTTGCCCACGCCAAGCCCCACCACCAGTGAAGCCATCACGACCGTGAACTTGTCTGAG 1500

Qy   1501 AATAGCCAGTGTTGTTGTCATCCCCCACCCGACTACTTGGAAATGCTTGTCTCCATGGAGCT 1560
Db   1501 AATAGCCAGTGTTGTTGTCATCCCCCACCCGACTACTTGGAAATGCTTGTCTCCATGGAGCT 1560

Qy   1561 GCCGACAGGAGGACGCAGAGATTCGGCGACAGCCCAGGACACTCCACCTTTAAGGGCCCCAAG 1620
Db   1561 GCCGACAGGAGGACGCAGAGATTCGGCGACAGCCCAGGACACTCCACCTTTAAGGGCCCCAAG 1620

Qy   1621 ACCCTGCCCACGCCCAAGCCCCACCAGTGACGTGCAAAAGCCTGACCACTTTTCTCAGCGGA 1680
Db   1621 ACCCTGCCCACGCCCAAGCCCCACCAGTGACGTGCAAAAGCCTGACCACTTTTCTCAGCGGA 1680

Qy   1681 GAAAACCCCTTTGCCACTGTGAAACTCCCCCCGACTGTGACGAATGACTGGCTCAGCACCC 1740
Db   1681 GAAAACCCCTTTGCCACTGTGAAACTCCCCCCGACTGTGACGAATGACTGGCTCAGCACCC 1740

Qy   1741 ATCATTCGATGAGAGGACGAGCCAAGCACTCTCCCGGGCCTCTCCCGTTCTCCCTTGCGGA 1800
Db   1741 ATCATTCGATGAGAGGACGAGCCAAGCACTCTCCCGGGCCTCTCCCGTTCTCCCTTGCGGA 1800

Qy   1801 ATGATGGGGGCGCATCCTGTCTGCACGGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860
Db   1801 ATGATGGGGGCGCATCCTGTCTGCACGGTGCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCT 1860

Qy   1861 AACTCCTAAATGTCGCCTTTAAGCAGAAAATCATGCTTCTCTCGTTTCACGTAGTTGGCGTTGAC 1920
Db   1861 AACTCCTAAATGTCGCCTTTAAGCAGAAAATCATGCTTCTCTCGTTTCACGTAGTTGGCGTTGAC 1920

Qy   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCCATTCAATCTTTA 1980
Db   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCCATTCAATCTTTA 1980

Qy   1981 ATTTTCTTCCTTATTCTGTTCAAGAAAAACAGTAAAACTTGGTTTCAATCTTTTA 2030
Db   1981 ATTTTCTTCCTTATTCTGTTCAAGAAAAACAGTAAAACTTGGTTTCAATCTTTTA 2030
```

```
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Zhiwei
;  TITLE OF INVENTION: No. US2004008249A1el Nucleic Acids and
;  TITLE OF INVENTION:  Secreted Polypeptides
;  FILE REFERENCE: 805A
;  CURRENT APPLICATION NUMBER: US/10/112,944
;  CURRENT FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 924
;  SOFTWARE: pt_FL_genes Version 5.0
;  SEQ ID NO 39
;  LENGTH: 2471
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (61)..(1251)
US-10-112-944-39
```

```
Query Match      72.3%;  Score 1514.6;  DB 7;  Length 2471;
Best Local Similarity   99.7%;  Pred. No. 1.9e-306;
Matches 1528;  Conservative   0;  Mismatches   4;  Indels   1;  Gaps   1;
```

```
Qy   499 AAATTCCACAAAGAGAATTATCCATGAGCTGGAGAAGAAGAGAATAGAACTTGACGTGAAATAT 558
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AAAATTCCACAAAGAGAATTATTCCATGAGCTGGAGAAGAAGAGAATAGAACTTGACGTGAAATAT 60

Qy   559 ATGAACCGAACTCTAAAAAGATACCAAAACAGAGACACAAGAGAAGAGAATAAATTAGAGTCTTTGGAG 618
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATGAACCGAACTCTAAAAAGATACCAAAACAGAGAATAAATTAGAGTCTTTGGAG 120

Qy   619 AAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAACCGAAACCACTC 678
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAGCCAAGGAAACCGAAACCACTC 180

Qy   679 AAATATGAACACAACAAAGAAAATTGAGTATGTGGAGACCGTTACTTCGTCAGAGTGAAATC 738
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAATATGAACACAACAAAGAAAATTGAGTATGTGGAGACCGTTACTTCGTCAGAGTGAAATC 240

Qy   739 CAGAAAATTCATTCGAGATGGTTGCAAAGAGGCTCTGCTTGACGAAGAGGCGCCTTCTGC 798
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGAAAATTCATTCGAGATGGTTGCAAAGAGGCTCTGCTTGACGAAGAGGCGCCTTCTGC 300

Qy   799 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAAACACATACATTATTATCACTTACTTACAGTCT 858
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TTTCTGGTTGATAAGCACTGTGGCTTTGCAAAACACATACATTATTATCACTTACTTACAGTCT 360

Qy   859 GCAGAACTACTGAATTCCAAGCTGCCTCCGGTGGCCAGGAGAACCTGTGTTGAATGCCATCAAA 918
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCAGAACTACTGAATTCCAAGCTGCCTCCGGTGGCCAGGAGGACCTGTGTTGAATGCCATCAAA 420
```

```
Db   361 GCAGAACTACTGAATTCCAAGCTGCCTCCGGTGGCCAGGAGGACCTGTGTTGAATGCCATCAAA 420
Db   421 GTGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGAACCCCAGCCTTCTACCCCGTG 480
Db   481 TCTCGAACTCCTCAGGCTTCACCCATGATCGAGAGAGAAGCAAATGTGGTTAGGGAAGATTAC 540
Db   541 GACACCCGTTTCTAAAATGCTCACCAAAGATGCCCCCCGGCTCCTTCAGGCGAGCCATACC 600
Db   601 AGTCCCTTGATCGATTATGTTTAATTAACCAGCACAGGGCTGCCCCGGAAATTCACAAAGGGTA 660
Db   661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACACGGATCAGTTTCGGTTGCAACG 720
Db   721 GGACTGAACATGATGAAGAAGCAGAAAGTGAAGAACCATCTTCCCGCCACACTGCGGGCTCC 780
Db   781 AACAAGAACCTTACTCACGTTTGCACAGGAGATGTCATCACGTTCATCCCCGAGGAG 840
Db   841 AAGGATGGCTGGCTCTTATGGAGAACACAGACGTGTGCCAAGGCGAGGGGTTGGTTCCCGTCG 900
Db   901 TCGTTACACGGAAGTTGCTGGAGAAAAATGAGACAGTGACCGGTGCCACCGCCAAGC 960
Db   961 CCCACACCAGTGAGGAAGCATCAGCACCCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATC 1020
Db  1021 CCCCCCACCCGACTACTTGGAATGCTTATCCATGGGGCAGCTGCCGACAGGAGACCAGAT 1080
Db  1081 TCGGTCAGGACGACCACATCCACACTTTAAGGCCCCAGTGTTCAAGGCCCGAGACGCGGCTCC 1140
Db  1141 AACGGATGCCAAACGGGACGTCGCAAAGCCGCCTTTTCTCAGCGGAGAAAAACCCCTTGCCACT 1200
Db  1201 GTGAAAACTCCGCCCGACTCGTGACCGAAATGATCGCTCGGCAACCATCATTCGGATGACGAGGAC 1260
Db  1261 AGCCAAGGAGGACTCTCCCGGGCCTCTCCGGTTCTCCGGGAAATGGGCGCGCATCCTGT 1320
Db  1321 CTGCCACCGGTGCTGACGGTGGAGCTTCAGTGAGACGCCTACTCTAATGTTCCCGC 1380
Db  1381 TTAAGCAAAATCATGCCTTCTCTGTTTCACGGTAGTTGGGTTGACAAGTTTCTGCCTTTAAG 1440
Db  1441 ATAAATGACTAATTAGTCTTAATGACCAGCTCAGCCATTTAAAAATATTTTTCTTCCTATTCTG 1500
```

```
Qy    1998  TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 2030
            ||||||||||||||||||||||||||||||||
Db    1501  TTCAAGAAACAGTAAACTTGGTTTCAATCTTTA 1533
```

RESULT 4

US-10-112-944-523
; Sequence 523, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1e1 Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 523:
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-112-944-523

Query Match   69.1%; Score 1448.4; DB 7; Length 1510;
Best Local Similarity   99.2%; Pred. No. 1.1e-292;
Matches 1498; Conservative 0; Mismatches 6; Indels 6; Gaps 4;

```
Qy    217  ATGTCCCGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAACTGTTATG  276
           ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      1  ATGTCCCGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGAAATGTTATG   60

Qy    277  GAACAGTTCAATCCTGGGCTCGGAAATTTAATAAACCTGGGAGAGAAAGCT  336
           |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     61  GAACAGTTCAATCCTGGGCTCGGAAATTTAATAAACCTGGGAGAGAAAGCT  120

Qy    337  GTAAACGCTATTGATCCTCGGCAGGAAAAAGCCTACTACCGATGGCGTGAG  396
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTAAACGCTATTGATCCTCGGCAGGAAAAAGCCTACTACCGATGGCGTGAG  180

Qy    397  ATTGCCACTCGGGTCCCCGTGTCAAC-TGAACTGGGACATGTCCTCATAG  455
           |||||||||||||||||||||||||| |||||||||||||||||||||||
Db    181  ATTGCCACTCGGGTCCCCGTGTCAACTTGAACTGGGACATGTCCTCATAG  240

Qy    456  TACCCACAAGAAAC-TCAACGAGAGTCTT--GATGAAAATTTAAAAAATTCCACAAAG  511
           |||||||||||||| |||||||||||||   |||||||||||||||||||||||||
Db    241  TACCCACAAGAAACTTCAACGAGAGTCTTCTGATGAAAATTTAAAAAATTCCACAAAG  300

Db    301  AGATTATCCATTGACGCTGGAGAAGAAGAGATAGAACTTCTTTGATGAAGAACTT  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    512  AGATTATCCATTGACGCTGGAGAAGAAGAGATAGAACTTCTTTGATGAAGAACTT  570

Db    361  CTAAAAAGATACCAAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    571  CTAAAAAGATACCAAAACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCT  630

Db    421  GAGTTGAAGAAGATCAGAAAGCCAAGGAAGACCCAGAAAACGCACTCAAATATGAAACAC  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    631  GAGTTGAAGAAGATCAGAAAGCCAAGGAAGACCCAGAAAACGCACTCAAATATGAAACAC  690

Db    481  AAAGAAAATTGAGTATGTGGAGACCGGTTACTTCTCGTCACGAGTGAAATCCAGAAATTCATT  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    691  AAAGAAAATTGAGTATGTGGAGACCGGTTACTTCTCGTCACGAGTGAAATCCAGAAATTCATT  750

Db    541  GCAGATGGTTGCAAAGAGGCTTCGCTTGAAGAGAAGAGAGCGCTTCTGCTTTCTGGTTGAT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    751  GCAGATGGTTGCAAAGAGGCTTCGCTTGAAGAGAAGAGAGCGCTTCTGCTTTCTGGTTGAT  810

Qy    811  AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG  870
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  AAGCACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTG  660

Qy    871  AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAAA  930
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAGTGCCAGAAA  720

Qy    931  ATCATGAATATGATCGAAGAAAAATAAAAGACCCCAGCCTCTTACCCCGTGTGTCTGGAACTCCT  990
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ATCATGAATATGATCGAAGAAAAATAAAAGACCCCAGCCTCTTACCCCGTGTGTCTGGAACTCCT  780

Qy    991  CAGGCTTCACCCATGATCGAGAGAGAAGCAATGTGGTTAGGAGAAAGATTAGGACACACCTTTCT 1050
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CAGGCTTCACCCATGATCGAGAGAGAAGCAATGTGGTTAGGAGAAAGATTAGGACACACCTTTCT  840

Qy   1051  AAATGCTCACCAAAGATGCCCCCGCTCCTTCAGGCAGAGCAATATACCAGTCCAGTCCCTTGATC 1110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  AAATGCTCACCAAAGATGCCCCCGCTCCTTCAGGCAGAGCAATATACCAGTCCAGTCCCTTGATC  900

Qy   1111  GATATGTTAATAACCCAGCCACGGCTCGCCCCGAATTCACAAAGGGTAAATAATAATTCAACA 1170
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  GATATGTTAATAACCCAGCCACGGCTCGCCCCGAATTCACAAAGGGTAAATAATAATTCAACA  960

Qy   1171  GGTACTTCCGACGATCCCAGTTTACACGGATCAGTTTCGGTTGCAACCGGACTTGAACATG 1230
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GGTACTTCCGACGATCCCAGTTTACACGGATCAGTTTCGGTTGCAACCGGACTTGAACATG 1020

Qy   1231  ATGAAGAAGCAGAAAAGTGAAGACCATCTTCCCGCACACACTGCGGCTCCAACAAGACCTTA 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ATGAAGAAGCAGAAAAGTGAAGACCATCTTCCCGCACACACTGCGGCTCCAACAAGACCTTA 1080

Qy   1291  CTCAGCTTTGCACAGGGAGAGATGTCATCACGGTGCTGCTCATCCCCGAGGAAGGAATGGCTGG 1350
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CTCAGCTTTGCACAGGGAGAGATGTCATCACGGTGCTGCTCATCCCCGAGGAAGGAATGGCTGG 1140

Qy   1351  CTCTATGGAGAACACGACGTGTCCAAGCGCGAGGGGGTTGGTTCCCGTTGCTGTTGTTCACCGAG 1410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CTCTATGGAGAACACGACGTGTCCAAGCGCGAGGGGGTTGGTTCCCGTTGCTGTTGTTCACCGAG 1200

Qy   1411  TTGCTTGAAGAAAATGAGACACAGAAAGCAGTTGACCGGTGCCAGCCCAAGCCACACACAGTG 1470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TTGCTTGAAGAAAATGAGACACAGAAAGCAGTTGACCGGTGCCAGCCCAAGCCACACACAGTG 1260

Qy   1471  AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCGAC 1530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCGAC 1320

Qy   1531  TACTTGGAATGCTTGTCCATGGGGCAGCTGCCGACAGGAGAGCAGAATTCGGCCAGGACG 1590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

Db 1321 TACTTGGAATGCTTATCCATTGGGGGGGGCAGCTGCCAGAGATTCGGCAGGACG 1380
Qy 1591 ACATCCACCTTTAAGGCCCCCAAGGGTCCAAGCGCCGAGACCGGGCTTCTAAGGATGCCAAC 1650
Db 1381 ACATCCACCTTTAAGGCCCCCCAGGGTCCAAGCGCCGGGCTTCCTAAGGATGCCAAC 1440
Qy 1651 GGGGACTGCAAAGCCCCCTTTTCTCACGCGGAGGAAAAACCCCTTTGCCACTGTGAAACTCCGC 1710
Db 1441 GGGGACTGCAAAGCCCGCCCTTTTCTCACGCGGAGGAAAAACCCCTTGCCACTGTGAAACTCCGC 1500
Qy 1711 CCGGACTGTGA 1720
Db 1501 CCGGACTGTGA 1510

RESULT 5
US-10-172-118-1805
Sequence 1805, Application US/10172118
Publication No. US20030224374A1
GENERAL INFORMATION:
  APPLICANT: Dai, Hongyue
  APPLICANT: He, Yudong
  APPLICANT: Linsley, Peter
  APPLICANT: Mao, Mao
  APPLICANT: Roberts, Chris
  APPLICANT: Van 't Veer, Laura
  APPLICANT: Van de Vijver, Marc
  APPLICANT: Bernards, Rene
  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
  FILE REFERENCE: 9301-175-999
  CURRENT APPLICATION NUMBER: US/10/172,118
  CURRENT FILING DATE: 2002-06-14
  PRIOR APPLICATION NUMBER: 60/380,770
  PRIOR FILING DATE: 2002-05-14
  NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1805
  LENGTH: 2120
  TYPE: DNA
  ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
  DATABASE ACCESSION NUMBER: NM_018842
  DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1805

Query Match   58.0%; Score 1216.4; DB 6; Length 2120;
Best Local Similarity   98.3%;
Matches 1229; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy   232 GAGGAGGTGAACCCGGCCTCACGGAGGAGCACCTACCGGAATGTTATGGAACAGTCAATCCT 291
Db   292 GTGGGGGGTGAAACTCCCCCTGTAATTAGGATCGCCGAGAATGTTATGGAACAGTTCAATCT 351
Qy   292 GGGGCTGCCAAATTAATAAACCTGGGGAAAAATTATTGACAAACGCTATGGCTATGATC 351
Db   352 GGGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC 411
Qy   352 CTGGCAGGAGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGAGATTGGCCACTGGGGTCC 411
Db   412 CTGGCAGGAGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGGCCACTGGGGTCC 471
Qy   412 CCCCTGTCAACTGAACTGGGACATGTCCTCCATAGAGAGATTTCAAGTAACCCCACAAGAAACTC 471
Db   472 'CCCGTGTCAACTGAACTGGGACATGTCCTTCCATAGAGAGATTTCAAGTAACCCCACAAGAAACTC 531
Qy   472 AACGAGAGTCTTGATGAAAAATTTTAAAAAAATTTCCACAAACAGATTATTCCATCATGAGCTGGAG 531
Db   532 AACGAGAGTCTTGATGAAAAATTTTAAAAAAATTCCACAAACAGATTATTCCATGAGCTGGAG 591
Qy   532 AAGAGAATAGAACTTGACGTGAAAATATATGCAACTCTTAAAAAGATATACCAAACAGAA 591
Db   592 AAGAAGATAGAACTTGACGTGAAAATATATGCAACTCTTAAAAAGATATACCAAACAGAA 651
Qy   592 CACAAGAATAAATTAGAGTCTTTGGAGAAAATCCCAAGCTGAGTTGAAGAAGAAGATCAGAAGG 651

Db   652 CACAAGAATAAATTAGAGTCTTTGGAGAAAATCCCAAGCTGAGTTGAAGAAGAAGATCAGAAGG 711
Qy   652 AAAAGCCAAGGAAGCCGAAACGGCACTCAAAATATGAACACAAAAGAAAATTGAGTATGTCGAG 711
Db   712 AAAAGCCAAGGAAGCCGAAACGGCACTCAAAATATGAACACAAAAGAAAATTGAGTATGTCGAG 771
Qy   712 ACCGTTACTTCTCGTCAGAGTGAAATCAGAAAATTCAGAGAATTCAGTGCAGAAAAAGAGGCT 771
Db   772 ACCGTTACTTCTCGTCAGAGTGAAATCAGAAAATTCAGTGCAGAAAAAGAGGCT 831
Qy   832 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCCTTCCGTGG 891
Db   832 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCCTTCCGTGG 891
Qy   892 CAGGAGACCTGTGTTGTTGATGCCATCAAAGTGCCAGAGAAAATCATTGAATATGATCGGAAGAA 951
Db   892 CAGGAGACCTGTGTTGTTGATGCCATCAAAGTGCCAGAGAAAATCATTGAATATGATCGGAAGAA 951
Qy   952 ATAAAGACCCCTAACCCCCGTGTCTCAGGGGAAAACTCCTCAGGGCTTCACCCATGATCGAG 1011
Db   952 ATAAAGACCCCAGCCTCTTACCCCCGTGTCTCAGGGGAAAACTCCTCAGGGCTTCACCCATGATCGAG 1011
Qy 1012 AGAAGCAAATGGTGGTTAGGAAAGATTAGCGACACCCTTTCTAAAATGCTCACCAAAAGATGCCC 1071
Db 1012 ATAAAGACCCCAGCCTCTTACCCCCGTGTCTGGGAAACTCCTCAGGGCTTCACCCATGATCGAG 1071
Qy 1072 CCCGCTCCTTCAGGCAGCAGCATATACCAGTCCCTTGATCGATAGTTTAATAACCCAGCC 1131
Db 1072 AGAAGCAAATGGTGGTTAGGAAAGATTAGCGACACCCTTTCTAAAATGCTCACCAAAAGATGCCC 1131
Qy 1132 ACGGCTGCCCCGGAATTCACAAAGGGTAAATAATTCAACAGGGTACTTCCGGAAGATCCCAGT 1191
Db 1132 CCCGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATAGTTTAATAACCCAGCC 1191
Qy 1192 TTACAGCGGATCAGTTTCGGTTGCAAACGGGGACTGAACATGAAGAAGACGAGAAGTGAAG 1251
Db 1192 ACGGCTGCCCCGGAATTCACAAAGGGTAAATAATTCAACAGGGTACTTCCGGAAGATCCCAGT 1251
Qy 1252 ACCATCTTCCCGCACACTCGGGGCTTCCAACAAGACCTTACTCAGCTTGCACAGGGACAGAT 1311
Db 1252 TTACAGCGGATCAGTTTCGGTTGCAAACGGGGACTGAACATGAAGAAGACGAGAAGTGAAG 1311
Qy 1312 GTCATCACGCTGCTCATCCCCGAGGAGAAGGAATGGCTGGCCTCTTATGGGACGAACACGACCGTG 1371
Db 1312 ACCATCTTCCCGCACACTCGGGGCTTCCAACAAGACCTTACTCAGCTTGCACAGGGAGAT 1371
Qy 1372 TCCAAGGCGAGGGGGTTGGGTTCCCGTGCTACACGAAGTTGCTGGAGAAAATGAGACA 1431
Db 1372 GTCATCACGCTGCTCATCCCCGAGGAGAAGGAATGGCTGGCCTCTTATGGGACGAACACGACCGTG 1431
Qy 1432 GAAGCAGTGACCGTGCCCAGGCCAAGCCCCACCAGTGAGAAGCAACATCAG 1481
Db 1432 TCCAAGGCGAGGGGGTTGGGTTCCCGTGCTACACGAAGTTGCTGGAGAAAATGAGACA 1491
Db 1492 GAAGCAGTGACCGTGCCCAGGCCAAGCCCCACCAGTGAGAAGCTTCAG 1541

RESULT 6
US-10-342-887-1805
Sequence 1805, Application US/10342887
Publication No. US20040058340A1
GENERAL INFORMATION:
  APPLICANT: Dai, Hongyue
  APPLICANT: He, Yudong
  APPLICANT: Linsley, Peter S.
  APPLICANT: Mao, Mao
  APPLICANT: Roberts, Christopher J.
  APPLICANT: Van 't Veer, Laura Johanna
  APPLICANT: Van de Vijver, Marc J.
  APPLICANT: Bernards, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/380,710
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1805
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo.sapiens
US-10-342-887-1805

Query Match 58.0%; Score 1216.4; DB 7; Length 2120;
Best Local Similarity 98.3%; Pred. No. 4.4e-244;
Matches 1229; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy    232 GAGGAGGTGAACCGGCTCACCGGAGAGCACCTACCCGGAATGTTATGAACAGTTCAATCCT 291
          | ||||| | |                     ||||||||||||||||||||||||||||||
Db    292 GTGGGGGGTGAAACTCCCCTGTAAATTAGGGATCGGCAGATGTTATGAACAGTTCAATCCT 351

Qy    292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGGACTAATGTATGAGATC 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 GGGCTGCGAAATTTAATAAACCTGGGAAAAATTATGAGAAAGCTGTAAACGGACTAATGTATGAGATC 411

Qy    352 CTGGGAGGAAAAGCCTACTACGGATGGAGTGGCCAAGATCGGGTGAGATTGCCACTGGGTCC 411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 CTGGCAGGAAAAGCCTACTACGGATGGAGTGGCCAAGATCGGGTGAGATTGCCACTGGGTCC 471

Qy    412 CCCGTGTCAACTGAACTGGGACACATGTTCCTCATAGAGAATTTCAAGTACCCACAAGAAACTC 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 CCCGTGTCAACTGAACCTGGGACACATGTTCCTCATAGAGAATTTCAAGTACCCACAAGAAACTC 531

Qy    472 AACGAGAGTCTTGATGAATTTTAAAAAAATTCAAGAAGACTTATCCATGAGCTGGGAG 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 AACGAGAGTCTTGATGAAATTTTAAAAAAATTCCACAAGAAGACTTATCCATGAGCTGGGAG 591

Qy    532 AAGAGAGATAGAACTTGACGTGAAATATATGAAGCGAACCAAACCAAACAGAA 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 AAGAGAGATAGAACTTGACGTGAAATATATGAAACGAACCAAACCAAACAGAA 651

Qy    592 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTCAGTTGAAGAGAGATCAGAAGG 651
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 CACAAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTCAGTTGAAGAGAGATCAGAAGG 711

Qy    652 AAAAGCCAAGGAGGACGGCCGAAACGCACACTCAAATATGAACACAAAAGAAAATTGAGTAGTAGTGGAG 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    712 AAAAGCCAAGGAGGACGGCCGAAACGCACTCAAATATGAACACAAAAGAAAATTGAGTAGTAGTGGAG 771

Qy    712 ACCGTTACTTCTCCTCAGAGTGAAAATCCAGAAAATCATTGCAGATGGTTGCAAAGAGGCT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    772 ACCGTTACTTCTCCTCAGAGTGAAAATCCAGAAAATCATTGCAGATGGTTGCAAAGAGGCT 831

Qy    772 CTGCTTGAAGAGAGAGGCGCTTCTGCTTTCTCGGTTGATAAGCACTGTGGCTTTGCCAAAC 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    832 CTGCTTGAAGAGAGAGGCGCCTTCTGCTTTCTCGGTTGATAAGCACTGTGGCCTTTGCAAAC 891

Qy    832 CACATACAATTATTATTCACTTCACTCTGCAGAACTACTCTGAATTCCAAGCTGCCTCGGGTCG 891
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 CACATACAATTATTATTCACTTCACTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGGTCG 951

Qy    892 CAGGAGGACCTGTCGTTGATCCCATCAAAAGTGCCACAGAAAATCATGAATATGATCGGAAGAA 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    952 CAGGAGGACCTGTCGTTGATGCCATCAAAAGTGCCCATGAATATGATCGGAAGAA 1011

Qy    952 ATAAAGACCCCAGCCTCTACCCCCCGTGTCGTCGGAACTCCTCAGGCTTCACCCATGATCGAG 1011
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| /
Db   1012 ATAAAGACCCCAGCCTCTACCCCCCGTGTCTCGGAACTCCTCAGGCTTCACCCATGATCGAG 1071
```

RESULT 7
US-10-450-763-9761
Sequence 9761, Application US/10450763
Publication No. US2005019675A1.
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 9761
LENGTH: 1729
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (319)..(939)
OTHER INFORMATION: SIMILAR
OTHER INFORMATION: 67% homologous to Homo sapiens insulin receptor tyrosine
OTHER INFORMATION: kinase substrate, accession number AF119666, Smith-Waterman Score=
OTHER INFORMATION: 619.
US-10-450-763-9761

Query Match 51.1%; Score 1071; DB 9; Length 1729;
Best Local Similarity 95.1%; Pred. No. 1.1e-213;
Matches 1205; Conservative 0; Mismatches 45; Indels 17; Gaps 9;

```
Qy   1012 AGAAGCAATGTGGTTAGGAAAAGATTACGACACCCCTTTCTAAAATGTCTCACCAAAGATGCCC 1071
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1072 AGAAGCAATGTGGTTAGGAAAAGATTACGACACCCCTTTCTAAAATGTCTCACCAAAGATGCCC 1131

Qy   1072 CCCGGCTCCTTCAGGCAGAGCAATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1132 CCCGGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCC 1191

Qy   1132 ACGGGCTGCCCCCGAAATTCACCAAAGGGTAAATAATTCAACAGGTTACTTCCGAAGGATCCCAGT 1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1192 ACGGGCTGCCCCCGAAATTCACCAAAGGGTAAATAATTCAACAGGTACTTCCGAAGGATCCCAGT 1251

Qy   1192 TTACAGGCGGATCAGTTTCGGTTGCAACGGGACTGGACTGAAACATGATGAAGAAGCAGAAGAGTGAAG 1251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1252 TTACAGGCGGATCAGTTTCGGTTGCAACGGGACTGGACTGAAACATGATGAAGAAGCAGAAGAGTGAAG 1311

Qy   1252 ACCATCTTCCCCGCACACTGCGGCGGCTCCAACAACAGACCTTACTCACTCAGCTTTGCACAGGGGACGTG 1311
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1312 ACCATCTTCCCCGCACACTGCGGGGCTCCAACAACAGACCTTACTCAGCTTTGCACAGGGGAGAT 1371

Qy   1312 GTCATCACGGCTGCTCATCCCCCGAGGAGGAAGAGATGGGCTGGGCTTCATGTGGAGAACACGACGTG 1371
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1372 GTCATCACGGCTGCTCATCCCCCGAGGAGGAAGAGATGGGCTGGGCTTCATGTGGAGAACACGACGTG 1431

Qy   1372 TCCAAGGCGAGGGGGTTGGTTCCCGTTGCTCGTACACGGAAGAGTTGCTGGAGAAAATGAGACA 1431
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1432 TCCAAGGCGAGGGGGTTGGTTCCCGTGCTCGTACACGGAAGAGTTGCTGGAGAAAATGAGACA 1491

Qy   1432 GAAGCAGTGACGGTGCCCACGGCCAAGCCCCACACCAGTGAGAAGCATCAG 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1492 GAAGCAGTGACGGTGCCCACGGCCAAGCCCCACACCAGTGAGAAGCTTCAG 1541
```

Query Match 51.1%; Score 1071; DB 9; Length 1729;
Best Local Similarity 95.1%; Pred. No. 1.1e-213;
Matches 1205; Conservative 0; Mismatches 45; Indels 17; Gaps 9;
US-10-450-763-9761

```
Qy    232 GAGGAGGTGAACCGGCTCACCGGAGAGCACCTACCCGGAATGTTATGAACAGTTCAATCCT 291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277 GTGGGGGGTGAAACTCCCTGTAATTAGGATCGCCAGAAGTTATGGAACAGTTCAATCCT 336

Qy    292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGGCTAATGATC 351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   337  GGGCTGCGGAAATTTAATAAACCTGGGGAAAAATTATGAGAAAGCTGTAAACGCTATGATC  396

Qy   352  CTGGCGAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGGCACTCGGTCC  411
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   397  CTGGCAGGAAAAGCCTACTACGATGGAGTGGCCAAGATCGGTGAGATTGGCACTCGGTCC   456

Qy   412  CCCGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTTCAAGTACCCACACAGAAAC-  469
          |||||||||||  |||||||||||||||||||||||||||| ||||||||||||||||||
Db   457  CCCGTGTCAACTTGAACTGGGACATGTCCTCATAGAGATTTCAGTACCCACACAGAAACT   516

Qy   470  TCAACGAGAGTCTT---GATGAAAATTTTAAAAAATTCACAAAGAGATTATCCATGAGC    526
          |||||||||||||    ||||  ||||||| |||||||||||||||||||||||||||||
Db   517  TCAACGAGAGTCTTTGATGGAAAATTTTTAAAAAATTCACAAAGAGATTATCCATGAGC    576

Qy   527  TGGAGAAGAAGAATAGAAACTT-GAGCTGAAATATATAATGAACGCAACTCTTAAAAAGATACCAA  585
          ||||||||||| ||||||||||  | | ||||| |||| | ||||||||||||||||||||||||
Db   577  TGGAGAAGAAGATAGAAACTTGGACCGTGAAAATATATCGAACGCAACTCTTAAAAAGATACCAA  636

Qy   586  ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCAAGCTGAAGTTGAAGAAGATC   645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   637  ACAGAACACAAGAATAAATTAGAGTCTTTGGAGAAATCCAAGCTGAAGTTGAAGAAGATC   696

Qy   646  AGAAGGAAAAAGCCAAGGAAGCCCGAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT  705
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   697  AGAAGGAAAAAGCCAAGGAAGCCGAAAACGCACTCAAATATGAACACAAAGAAATTGAGTAT  756

Qy   706  GTGGAGAACCGTTACTTCTCGTCAGAGTGAAATCC-AGAAAATTCATTGCAGATTATTGCAA  764
          ||||||||||||||||||||||||||||||||||| |||||| |||||||||  | |||||
Db   757  GTGGAGAACCGTTACTTCTCGTCAGAGTGAAATCCTAGAAAATCCATTGCAAGANGGTTGCAA  816

Qy   765  AGAGGCTCTGCTTGAAGAGAAGAGAGGG-------CTTCTGCTTTCTGGTTGATAAGCACTG  818
          |||||||||||||||||||||||||||        ||||||||||||||||||||||||||
Db   817  AGAGGCTCTGCTTGAAGAGAAGAGAGGGCTTCTGCTTTCTGGTTGATAAGCACTG        876

Qy   819  TGGCTTTGCAAA---CCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATT-C  875
          ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||| |
Db   877  TGGCTTTGCAAAACCCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC   936

Qy   876  CAAGCTGCCTCGGTCGGCGAGCAGCCTCGTTCGTTGATGCCATCAAAGTG--CCAGAAGAAATCA  934
          ||||||||||||||||| |  | |||| | |||||||||||||||||||   |||||| ||||||
Db   937  CAAGCTGCCTCGGTCGGCAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCA          996

Qy   935  TGAATATGATCGAAGAAATAAAAGACCCCAGCCTCTACCCCGTGTCTGGAACTCCTCAGG   994
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   997  TGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCGTGTCTGGAACTCCTCAGG  1056

Qy   995  CTTCACCCATGATCGAGAGAAGCAAATGTGGTTAGGTAAGGAAAGAGATATAACGACACTG  1054
          |||||||| ||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  1057  CTTCACCCCATGATCGAGAGAAGCAAATGTGGTTAGGTAAGGAAAGAGATATACGACACTG  1116

Qy  1055  GCTCACCAAAAGATGCCCCCCCGCTCCTTCAGCGGCGAGCAGCATATACCAGTCCCTTGATCGATA  1114
          |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  1117  GCTCACCAAAGATGCCCCCCGCTCCTTCAGCGGCGAGCAGCATATACCAGTCCCTTGATCGATA   1176

Qy  1115  TGTTTAATAACCCAGCCACCAGAAATTCACAAAGGGTAAAATAATTCAACAGGTA        1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1177  TGTTTAATAACCCAGCCACCAGAAATTCACAAAGGGTAAAATAATTCAACAGGTA        1236

Qy  1175  CTTCCGAAGATCCCAGTTTACNCGCGATCCAGTTTCGGGTTGCAAGGGACTGAACATGATGA  1234
          |||||| ||||||||||||||||| |||||| ||||||||||||||||||||||||||||||
Db  1237  CTTCCGNAGATCCCAGTTTACNCGGCGATCAGTTTCGGGTTGCAAGGGACTGAACATGATGA  1296

Qy  1235  AGAAGCAGAAAAGTGAAGACACCATCTTCCCGCACACTGGGCGGCTTCCAACAAGACCTTACTCA  1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1297  AGAAGCAGAAAAGTGAAGACACCATCTTCCCGCACACTGGGCGGCTTCCAACAAGACCTTACTCA  1356

Qy  1295  GCTTTGCACAGGGGGAGTGTCATCACGCTGTGCTCATCCCGAGGAGGAGGAATGGCTGGGCTCT  1354
          |||||||||||||| ||| |||||||||||||||||||||||||||||| |||||||||||||||
Db  1357  GCTTTGCACAGGGGAGATGTCATCACGCTGTGCTCATCCCGAGGAGGAAGGAATGGCTGGGCTCT  1416

Qy  1355  ATGGAGAACACGAGCGTGTTCCAAGGCGAGGCGAGGGGTTGGTTCCCGTGCTGCGTGCTCGTTGC  1414
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  1417  ATGGAGAACACGAGCGTGTTCCAAGGCGAGGCGAGGGGGTTGGTTCCCGTGCTGCGTGCTCGTTGC  1476

Qy  1415  TGGAGAAGAATGAGACAGAAGAGAAGCAGTGACAGTGACCGTGCCCAAGGCCCCACAGTTGAGGA  1474
          |||| ||||| |||||||||| ||||||| ||||| ||||||||||||| |||||| ||||||| |
Db  1477  TGGAAGAAAATGAGACAGAAGCAGTGACCAGTGACCGTGCCCAAGCCCACACCAGTTGAGA      1536

Qy  1475  GCATCAG  1481
          || |||||
Db  1537  GCTTCAG  1543
```

```
Query Match             48.7%;  Score 1021.2;  DB 9;  Length 1993;
Best Local Similarity   89.3%;  Pred. No. 3e-203;
Matches 1211;  Conservative    0;  Mismatches   38;  Indels  107;  Gaps    6;

Qy   680  AATATGAACACAAAGAAATTGAGTATGTGGAGAGGACCGTTACTTCGTCGTGAGTGAAATCC  739
          || || |||   ||||||||||||| |||||||| |||||||||||| ||||||||||||||
Db   158  AACTTGAGGTGAAATATATGAACTATGTGGGAGAGACCGTTACTTCTCGTCGTGAGTGAAAATCC  217

Qy   740  AGAAAATTCATTCGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGAGGCGCGCTTCTCGCT  799
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||| |||
Db   218  AGAAATTCATTCGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGAGAGGCGCTTCGCT      277

Qy   800  TTCTGGTTGATAAGCACTGTGGCGCTTTGGCAAAACCACATACATTATTATCACTTACAGTCTG  859
          |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db   278  TTCTGGTTGATAAGCACTGTGGCTTTGGCAAAACCACATACATTATTATCACTTACAGTCTG     337

Qy   860  CAGAACTACTGAATTCCAAGCTCGCTCGGTGGCAGGAGACCTGGTTGATGCCATCAAAG  919
          |||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||||
Db   338  CAGAACTACTACTGAATTCCAAGCTCGCTCGGTGGCAGGAGACCTGTGTTGATGCCATCAAAG  397

Qy   920  TGCCAGAGGAAAATCATCGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCGTGT  979
          |||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   398  TGCCAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTCTACCCCGTGT     457

Qy   980  CTGGAACTCCTCAGGCTTCACCCCATGATCGAGAGAAGCAAATGTCGTTAGGAAAGAGATTACG  1039
          |||||||||||||||||||||||||||||||||||||||||||| |||    |
Db   458  CTGGAACTCCTCAGGCTTCACCCCATGATCGAGAGAAGCAAATGTG------              501

Qy  1040  ACACCCTTTCTAAATGCTCACCAAAGATGCCCCCGCTCCTTCAGGCAGGCATATATACCA  1099
```

Query Match   28.3%;   Score 594; DB 5; Length 625;
Best Local Similarity   100.0%; Pred. No. 4.3e-114;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 205 GCGGCCGCAGCCATGTCCCGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGGAGCACCTAC 264
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32 GCGGCCGCAGCCATGTCCCGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGGAGCACCTAC   91

Qy 265 CGGAATGTTATGGAACAGTTCAAATCCTGGGGCTGCAAAATTTAATAAACCTGGGGGAAAAT 324
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   92 CGGAATGTTATGGAACAGTTCAAATCCTGGGGCTGCAAAATTTAATAAACCTGGGGGAAAAT 151

Qy 325 TATGAGAAAGCTGTAAACCGTATGATCCTGGCAGAAAAAGCCTACTACGATTGGGAGTGGCC 384
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 152 TATGAGAAAGCTGTAAACCGTATGATCCTGGCAGAAAAAGCCTACTACGATTGGGAGTGGCC 211

Qy 385 AAGATCGGTGAGATTGCCACTGGGGTCCCCCGTCGTCAACTGAACTGGGACATGGTCCTCATA 444
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 212 AAGATCGGTGAGATTGCCACTGGGGTCCCCCGTCGTCAACTGAACTGGGACATGGTCCTCATA 271

Qy 445 GAGATTTCAAGTACCCACAAAGAAAACTCAACGAGAGAGTCTTGATGAAAATTTTAAAAATTC 504
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 272 GAGATTTCAAGTACCCACAAAGAAAACTCAACGAGAGAGTCTTGATGAAAATTTTAAAAATTC 331

Qy 505 CACAAAGAGATTATCCATGAGCTTGGAGAAGAAGATGAGACTTGACGTGAAATATATGAAC 564
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 332 CACAAAGAGATTATCCATGAGCTTGGAGAAGAAGATGAGACTTGACGTGAAATATATGAAC 391

Qy 565 GCAACTCTAAAAAGATACCAAACACAGAACACAAGAATAAAATTAGAGTCTTTGGAGAAATCC 624
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 392 GCAACTCTAAAAAGATACCAAACACAGAACACAAGAATAAAATTAGAGTCTTTGGAGAAATCC 451

Qy 625 CAAGCTGAGTTGAAGAAGATCAGAAGAAAAGCCGAAAACCGGACACTCAAATAT 684
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 452 CAAGCTGAGTTGAAGAAGATCAGAAGAAAAGCCGAAAACCGGACACTCAAATAT 511

Qy 685 GAACACAAAAGAAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAA 744
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 512 GAACACAAAAGAAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAA 571

Qy 745 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAGAAGAAGAGAGGCGCTTCTGC 798
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 572 TTCATTGCAGATGGTTGCAAAGAGGCTCTGCTTGAAGAGAGAAGAAGAGAGGCGCTTCTGC 625

Db   502 --------------------------------TCCCTGTTCCTTCTCGG------|| 517
                                          ||| | |||||| |||
Qy 1100 GTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCTGCCCGAATTCACAAAGGGTAA 1159

Db   518 ------------------GCAATGATCAAAACGAGG 535
                            ||||||||||||||||||
Qy 1160 ATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGAATCAGTTTCGGTTGCAACGG 1219

Db   536 AAATAGGACAGGGTACTTCCGAAGATCCCAGTTTACAGCGAATCAGTTTCGGTTGCAACGG 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1220 GACTGAACAATGATGAGAAGAAGCAGAGAAAAGTGAAGAACCAATCTTCCCGGACACTGCGCAGCTCCA 1279

Db   596 GACTGAACAATGATGAGAAGAAGCAGAGAAAAGTGAAGAACCAATCTTCCCGGACACTGCGCAGCTCCA 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1280 ACAAGACCTTACTCAGCTTTGCACAGGGACGATGTCATCACGGTCTCATCCCCGGCGGGGCTCCCA 1339

Db   656 ACAAGACCTTACTCAGCTTTGCACAGGGACGATGTCATCACGGTCTCATCCCCGGCGGGGCTCCCA 715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1340 AGGATGGGCTGGCTCTCTATGGAGAACGACGACGTGTCCAAGGGCGAGGGGGGGGTTGGTTCCCGTCGT 1399

Db   716 AGGATGGGCTGGCTCTCTATGGAGAACGACGACGTGTCCAAGGGCGAGGGGGGGGTTGGTTCCCGTCGT 775
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1400 CGTACACGGAAGTTGCTGGGAAGAAAAATGAGACAGCAGTGACCGGTGCCCACGCCAAGCC 1459

Db   776 CGTACACGGAAGTTGCTGGGAAGAAAAATGAGACAGCAGTGACCGGTGCCCACGCCAAGCC 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1460 CCACACCAGTGAGGAAGCATCAGCACCGTGAAACTTGTCTGTCGTGGTTGTTGTTCATCC 1519

Db   836 CCACACCAGTGAGGAAGCATCAGCACCGTGAAACTTGTCTGTCGTGGTTGTTGTTCATCC 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1520 CCCCACACCCGGACTACTTGTGAATGCTTGTCTGCTGCGACGGGGACAGGAGAGGCAGAGATT 1579

Db   896 CCCCACACCCGGACTACTTGTGAATGCTTGTCTGCTGCGACGGGGACAGGAGAGGCAGAGATT 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1580 CGGCCAGGACGACGCATCCACCTTTAAGGCCCCAGCCCCCAGAGCGGTCCTCCTA 1639

Db   956 CGGCCAGGACGACGTCCAAGCCTTTAAGGCCCCAGCCCCCAGAGCGGGTCCTCCTA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1640 ACGATGCCAAACGGGACTGCAGAACCCGCTTTCTCAGCGGAGAAAAACCCCTTTGCCACTG 1699

Db 1016 ACGATGCCAAACGGGACTGCAGAACCCGCTTTCTCAGCGGAGAAAAACCCCTTTGCCACTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1700 TGAAACTTCCGCCCCGACTGTGACGAATGATCGCTCGGCGGCACCCATCATTCGATGAGAGGACA 1759

Db 1076 TGAAACTTCCGCCCCGACTGTGACGAATGATCGCTCGGCGGCACCCATCATTCGATGAGAGGACA 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1760 GCCAAGGACTCTCCC--GGGCCTCCGGTTCTCCCTGGGAATGAGATGGGC--GCATCCT 1816

Db 1136 GCCAAGGACTCTCCCGGGGCCTCCGGTTCTCCCTGGGAATGAGATGGGCCGCCATCCT 1195
          |||||||||||||| |||||||||||||||||||||||||||||| |||||||
Qy 1817 GTCTGCCACGTGCTGGACGGTC-GGGAAGCTTCAGTGGAGGGCCTAACTCTAAATGGGAC 1875

Db 1196 GTCTGCCACGTGCTGGACGGTTCGGGAAGCTTCAGTGGAGGGCCTAACTCTAAATGGGAC 1255
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy 1876 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTCGAC-AAGTTTCTGCCCTT 1934

Db 1256 TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAAGTTTCTGCCCTTT 1315
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Qy 1935 AAGATAAATGACTAATAGTCTAAGCTCAGCCATTTAAAAATATTTCTTCCTATT 1994

Db 1316 AAGATAAATGACTAATAGTCTAAGCTCAGCCATTTAAAAATATTTCTTCCTATT 1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1995 CTGTTCAAGAAAACAGTAAAACTTGGTTTCAAATCTTTA 2030

Db 1376 CTGTTCAAGAAAACAGTAAAACTTGGTTTCAAATCTTTA 1411

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 269
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (46)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (46)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (460)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1233)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-269

Query Match   27.5%;   Score 576.4;   DB 3;   Length 1242;
Best Local Similarity   95.3%;   Pred. No. 2.8e-110;
Matches 592;   Conservative 1;   Mismatches 28;   Indels 0;   Gaps 0;

```
Qy  1410 GTTGCTCGGAAGAACAGAAATAGAGACAGAGCAGTGACCGTGCCCACGCCCACACCAGT 1469
         ||||||  |                                     |||||||||||||||
Db   239 GGTGGGTGAGACAATTATCTAACACGTGCTGTGTCCTTTTTGGGTTTGAGCCCACACCAGT 298

Qy  1470 GAGAAGCATCAGCACCGTGAACTTGTCTCGAGAATAGCAGTGTTGTCATCCCCCACCCGA 1529
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GAGAAGCATCAGCACCGTGAACTTGTCTCGAGAATAGCAGTGTTGTCATCCCCCACCCGA 358

Qy  1530 CTACTTGGAAATGCTTGTCCATGGGGGGCCGACAGCTGCCGACAGGAGGACCAGATTCGG 1589
         |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   359 CTACTTGGAAATGCTTATCCATGGGGGGCCACGTGCCCACAGGAGACAGATTCGGCCGCCAGGAC 418

Qy  1590 GACATCCACCTTTAAGGCCCCCCAGCGCTCCAAGCCCGAGACCCGCGGCTCCTCAACGATGCCAA 1649
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 GACATCCACCTTTAAGGCCCCCAGCGCTCCAAGCCCGAGACCCGNGGCTCCTCAACGATGCCAA 478

Qy  1650 CGGGACTGCAAAAGCCGCCCTTTTCTCAGCGGAGAAAAAACTCCG 1709
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   479 CGGGACTGCAAAAGCCGCCCTTTTCTCAGCGGAGAAAAAACCCCTTTGCCACTGTGAAACTCCG 538

Qy  1710 CCCGACTGTGACGAATGATCGGCTCGGCACCCCATCATTCGATGAGAGGACCAGCCAAGGACT 1769
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 CCCGACTGTGACGAATGATCGGCTCGGCACCCCATCATTCGATGAGAGGACAGCCAAGGACT 598

Qy  1770 CTCCCGGGCCTCCCGGTTCCTCCTTGCGGAATGATGGGCGCCATCCTGTCTGCTGCACGTGC 1829
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 CTCCCGGGCCTCCCGGTTCCTCCTTGCGGAATGATGGGCGCCATCCTGTCTGCCACGTGC 658

Qy  1830 TGACCGGTCGGGAAGCTTCACTGGAGAGAGGCCTAACTCTAATGTCGCCTCTTAAGCAAATC 1889
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TGACCGGTCGGGAAGCTTCACTGGAGAGAGGCCTAACTCTAATGTCGCCTCTTAAGCAAATC 718

Qy  1890 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTTTGCCTTTAAGATAAATGAGTAA 1949
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   719 ATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAGTTTCTTTGCCTTTAAGATAAATGAGTAA 778
```

RESULT 11
US-10-066-543-990
; Sequence 990, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indriias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5, 362, 424, 427, 444, 449, 565, 567, 584, 602, 609, 611,
; LOCATION: 624, 629
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-990

Query Match   27.5%;   Score 576.2;   DB 5;   Length 641;
Best Local Similarity   97.3%;   Pred. No. 2.3e-110;
Matches 578;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

```
Qy   205 GCGGCGCGCAGCCATGTCCCCGGGGGCCCGAGGGGTTGAAACCCGCTCACGGAGCACCTAC 264
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    48 GCGGCGCGCAGCCATGTCCCCGGGGGCCCGAGGGGTTGAAACCCGGCTCACGGAGCACCTAC 107

Qy   265 CGGAATGTTATGGAACAGTTCAATCCTGGGCGTCGGAAATTAATAAAACTTGGGGGAAAAT 324
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   108 CGGAATGTTATGGAACAGTTCAATCCTGGGCGTCGGAAATTTAATAAAACCTGGGGGAAAAT 167

Qy   325 TATGAGAAAGCTGTAAACGCTATGATCCTGCGCAGGAAAAGCCTACTGATGGATGGCC 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   168 TATGAGAAAGCTGTAAACGCTATGATCCTGGCGCAGGAAAAGCCTACTACGGATGGAGTGGCC 227

Qy   385 AAGATCCGGTGAGATTGCCACTGGGTCCCCGTGTCAAACTGAACTGGACACATGTCCTCATA 444
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   228 AAGATCCGGTGAGATTGCCACTGGGTCCCCGTGTCAAACTGAACTGGACACATGTCCTCATA 287

Qy   445 GAGATTTCAAGTACCCACACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 504
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   288 GAGATTTCAAGTACCCACACAAGAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAATTC 347

Qy   505 CACAAAGAGATTTATCCATGAGCTGGAGAAGAGATAGAACTTGACGTTGAAATATATGAAC 564
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   348 CACAAAGAGATTTATNCATGAGCTGGAGAAGAGATAGAACTTGACGTGAAAATATATGAAC 407

Qy   565 GCAACTCTAAAAAGATACCAAACAGAAACACAAGAAGAATAAAATTAGAGTGTCTTTGGAGAAAATCC 624
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAAGAATAAAATTAGAGTGTCTTTGGAGAAAATCC 624
```

(continuation of previous alignment — left column)

```
Db   408 GCAACTCTTAAAGATNCNNAACAGAACACAAGAGAATNAATTNGATGCTTTGGAGAAATCC 467
         ||||||||||||||||  || ||||||||||||||||  |||| |  ||||||||||||||
Qy   625 CAAGCTGAGTTGGAGAAGATCAGGAGGAAGATCAGGAAGCCAGAAACCGCACTCAAATAT 684
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   468 CAAGCTGAGTTGGAGAAGATCAGGAGGAAGATCAGGAAGCCAGAAACCGCACTCAAATAT 527
Qy   685 GAACACAAAAGAAATTGAGTATGTGGAGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAA 744
         ||||||||||||||||||||||||||||||||||||||||||  |||||||||||||| |||
Db   528 GAACACAAAAGAAATTGAGTATGTGGAGAGACCGTTACTTCTTNNGTCAGAGTGAAATCCANAAA 587
Qy   745 TTCATTGCAGATGGTTGCAAAAGAGGCTCTGCTTGAAGAAGGAGGACGGCGCTTCTGC 798
         ||||||||||||| |||||  |  |||||||||||||  || || |  |||||||||
Db   588 TTCATTGCAGATGGNTGCAAANANANGCTCTGCTTGAANAGAANAAAAGCGCTTCTGC 641
```

(continuation of previous alignment — right column)

```
Qy  1283 AGACCTTACTCAGCTTTGCACAGGGAGAGTGTCATCACGGCTGCTCATCCCGACGGAGAAGG 1342
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGACCTTACTCAGCTTTGCACAGGGAGAGTGTCATCACGGCTGCTCATCCCGACGGAGAAGG 480
Qy  1343 ATGGCTGGCTCTTATGGAGAACACGACGTGTCCAAGGCGCGAGGGGGTGGTTCCCGTCGTCGT 1402
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGGCTGGCTCTTATGGAGAACACGACGTGTCCAAGGCGCGAGGGGGTGGTTCCCGTCGTCGT 540
Qy  1403 ACACCGAAGGTTGCTCGGAAGA 1421
         ||||||||||||||||||||||
Db   541 ACACCGAAGGTTGCTCGGAAGA 559
```

RESULT 12

US-09-878-134-252

; Sequence 252, Application US/09878134
; Publication No. US20020863031A1
; GENERAL INFORMATION:
;   APPLICANT: Meagher, Madeleine Joy
;   APPLICANT: King, Gordon E.
;   APPLICANT: Xu, Jiangchun
;   APPLICANT: Secrist, Heather
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;   TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF COLON CANCER
;   FILE REFERENCE: 210121.532
;   CURRENT APPLICATION NUMBER: US/09/878,134
;   CURRENT FILING DATE: 2001-06-07
;   NUMBER OF SEQ ID NOS: 377
;   SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 252
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(559)
;   OTHER INFORMATION: n = A,T,C or G
US-09-878-134-252

Query Match   26.5%; Score 556; DB 3; Length 559;
Best Local Similarity   99.5%; Pred. No. 3.7e-106;
Matches 556; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy   863 AACTACTGAATTCCAAGCTGCTTCCGGTGGCAGAGAACCTGTTGTTGATCGTACCATCAAAGTGC 922
         |||||||||||||||||||||||| || |||||||| ||||||||||||||||||||||||||||
Db     1 AACTACTGAATTCCAAGCTGCTTCGGNNGGCAGAGACCTGTTGTTGATCGTACCATCAAAGTGC 60
Qy   923 CAGAGAAAATCATGAATAGATGAAGAGAAATAAAGACCCCAGCCTTCTTACCCCCGTGTCTG 982
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    61 CAGAGAAAATCATGAATAGATGAAGAGAAATAAAGACCCCAGCCTTCTACCCCCGTGTCTG 120
Qy   983 GAACTCCTCAGGCTTCACCCATATCAGTCGGTTAGGAAAAGAATTACGACA 1042
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAACTCCTCANNCTTCACCCATATCAGTCGGTTAGGAAAAGAATTACGACA 180
Qy  1043 CCCTTTCTAAAATGCTCACCAAAGATGCCCCCCCCCGCTCCTTCCTTCAGCCAGTC 1102
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCCTTTCTAAAATGCTCACCAAAGATGCCCCCCCCCGCTCCTTCCTTCAGCCAGTC 240
Qy  1103 CCTTGATCGATATGTTTAATAACCCAGCCACCACGGCTGCCCCGAATTCACAAAGGGTAAATA 1162
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTTGATCGATATGTTTAATAACCCAGCCACCACGGCTGCCCCGAATTCACAAAGGGTAAATA 300
Qy  1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATGGCGGCGGGAC 1222
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATGGCGGGGAC 360
Qy  1223 TGAACATGATGAAGAAGCAGAAGAGTGAAGACCATCTTCCGGCACACTGCGGGCTCCAACA 1282
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGAACATGATGAAGAAGCAGAAGAGTGAAGACCATCTTCCGGCACACTGCGGGCTCCAACA 420
```

RESULT 13

US-10-066-543-321

; Sequence 321, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;   APPLICANT: Jiang, Yuqiu
;   APPLICANT: Pyle, Ruth A.
;   APPLICANT: Xu, Jiangchun
;   APPLICANT: Indrias, Carol Yoseph
;   APPLICANT: Lodes, Michael J.
;   APPLICANT: Secrist, Heather
;   APPLICANT: Carter, Darrick
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Smith, Carole L.
;   APPLICANT: Durham, Margarita
;   APPLICANT: Stolk, John A.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
;   FILE REFERENCE: 210121.563
;   CURRENT APPLICATION NUMBER: US/10/066,543
;   CURRENT FILING DATE: 2002-01-31
;   NUMBER OF SEQ ID NOS: 3417
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 321
;   LENGTH: 539
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 20, 23, 51, 61, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-321

Query Match   24.8%; Score 520.8; DB 5; Length 539;
Best Local Similarity   98.7%; Pred. No. 8.4e-99;
Matches 522; Conservative   0; Mismatches   7; Indels   0; Gaps   0;

```
Qy   205 GCGGCCGCAGCCCATGTCCCGGGGCCCCGGAGGAGCGCTCACGGAGAGCACCTAC 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    11 GCGGCCGCANNCCNTGTCCCGGGGCCCCGGAGGAGGTGAACANGGTCACGGANAGCACCTAC 70
Qy   265 CGGAATGTTATGGAAACAGTTCAATCCTGGGCTCGGGAATTTAATTAAACCTGGGGGAAAAT 324
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71 NGGAATGTTATGGAAACAGTTCAATCCTGGGCTCGGGAATTTAATTAAACCTGGGGGAAAAT 130
Qy   325 TATGAGAAAGCTGTAAAGCCTATGATGATCCTGGAAGAAAAGCCTACTACGGATGGAGTGGGC 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 TATGAGAAAGCTGTAAAGCCTATGATGATCCTGGAAGAAAAGCCTACTACGGATGGAGTGGGC 190
Qy   385 AAGATCGGTGAGATTGCCAGTCCCCGTGTCCCGGGACATGGTCCTCATA 444
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 AAGATCGGTGAGATTGCCACTGGGGTCGCAGGGAAACTCAACGACACATGTCCCCATA 250
Qy   445 GAGATTTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGACGTGAAAAAAATTC 504
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 GAGATTTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGACGTGAAAATATGAAC 310
Qy   505 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAAACTTGACGTGAAAATATGAAC 564
         |||||||||||||||||||||||
Db   361 ...
```

Db   311 CACAAAGAGATTATCCATGAGCTGGAGAAGAAGATAGAACTTGACGTGAAATATATGAAC 370

Qy   565 GCAACTCTAAAAAGATTACCAAAACAGAACAGAAACAGAAGAATAAATTAGAGTCTTTG...GAAAATCC 624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 GCAACTCTAAAAAGATTACCAAAACAGAACAGAAACAGAAGAATAAATTAGAGTCTTTGGAGAAATCC 430

Qy   625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAAGCCAAGGAAACGCAAACTCAAATAT 684
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   431 CAAGCTGAGTTGAAGAAGATCAGAAGGAAAAAGCCAAGGAAACGCACTCAAATAT 490

Qy   685 GAACACAAAGAAAATTGAGTATGTGGGACACCGTTACTTCTTCGTCAGAGTG 733
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 GAACACAAAGAAAATTGAGTATGTGGGACACCGTTACTTCTTCGTCAGAGTG 539

RESULT 14
US-09-960-253-83
; Sequence 83, Application US/0996025
; Patent No. US2002123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-83

Query Match   23.3%; Score 488; DB 3; Length 511;
Best Local Similarity   99.6%; Pred. No. 6e-92;
Matches 510; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy   806 TTGATAAGCACTCTGGGCTTTGCAAAACCAACATACATTAATTATTACTTACAGTCTGCAGAAC 865
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTGATAAGCACTCTGGGCTTTGCAAAACCAACATACATTAATTATTACTTACAGTCTGCAGAAC   60

Qy   866 TACTGAATTCCAAGCTGCCTCGGTGGTGCGAGCGAGACGACCTGTGTTGTTGATGCCATCAAAAGTGCCAG 925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACTGAATTCCAAGCTGCCTCGGTGGTGCGAGCGAGACGACCTGTGTTGTTGATGCCATCAAAAGTGCCAG 120

Qy   926 AGAAAATCATGAATTGATCGAGAGAAAATAAAGACCCCAGCCTCTACCCCGTGTCTGGAA 985
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGAAAATCATGAATTGATCGAGAGAAAATAAAGACCCCAGCCTCTACCCCGTGTCTGGAA 180

Qy   986 CT-CCTCAGGCTTCACCCATGATCGAGAAGCAATGTGGTTAGGAGAAGAGATTACGACACC 1044
             || ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   181 CTCCCTCAGGCTTCACCCATGATCGAGAAGC--ATGTGGTTAGGAGAAGAGATTACGACACC 239

Qy   1045 CTTTCTAAATGCTCACCAAAGAATAGCCCCCGCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCC 1104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 CTTTCTAAATGCTCACCAAAGAATAGCCCCCGCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCC 299

Qy   1105 TTGATCGATATGTTTAATAACCCAGCCACCGGCTGCCCCCGAATTCACAAAGGGTAAAATAAT 1164
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 TTGATCGATATGTTTAATAACCCAGCCACCGGCTGCCCCCGAATTCACAAAGGGTAAAATAAT 359

Qy   1165 TCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTG 1224
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 TCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTG 419

Qy   1225 AACATGATGAAGAAGAGAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCCGGCGCTCCAACAAG 1284
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   420 AACATGATGAAGAAGAGAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCGCTCCAACACAAG 479

Qy   1285 ACCTTACTCAGCTTTGCACAGGAGGAGATGTCAT 1316
             ||||||||||||||||||||||||||||||||||
Db   480 ACCTTACTCAGCTTTGCACAGGAGGAGATGTCAT 511

RESULT 15
US-09-960-253-90
; Sequence 90, Application US/0996025
; Patent No. US2002123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-90

Query Match   22.4%; Score 470.4; DB 3; Length 512;
Best Local Similarity   99.8%; Pred. No. 2.9e-88;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 CTCCTCTGCTCTCCTTCGGAAGAGAGGCCGAGGGCGGGGGGGCCTGCCCGCAAGTTTTTGACATTTTCGCAG   60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 CTCCTCTGCTCTCCTTCGGAAGAGAGGCCGAGGGCGGGGGGGCCTGCCCGCAAGTTTTTGACATTTTCGCAG 100

Qy   61 CGGAGACGCGCGGCGGGGCCACTCTCCGGGCCCACCGGCCTGCGGCGGGGGCGGCGACCCCTTCAGAGC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 CGGAGACGCGCGGCGGGGCCACTCTCCGGGCCCACCGGCCTGCGGCGGGGGCGGCGACCCCTTCAGAGC 160

Qy   121 CCCTTAGTCGCGCGCCCCGGCCCCTCCCCGCTGGCCCGGAGTCCGGCGGCCGCCACGGAGGCCGCCGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 CCCTTAGTCGCGCGCCCCGGCCCCTCCCCGCTGGCCCGGAGTCCGGCGGCCGCCACGGAGGCCGCCGG 220

Qy   181 CGTCCTCCCCCGCGGCGCTTGCTCCCGCCCCCGGCGCCAGCGCCATGTCCCGGGGCGCCGAGGAGGTG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 CGTCCTCCCCCGCGGCGCTTGCTCCCGCCCCCGGCGCCAGCGCCATGTCCCGGGGCGCCGAGGAGGTG 280

Qy   241 AACCCGGCCTCACGGAGGCACTTACCGGGAATGTTATGGACACAGTTCAATCCTGGGGCTTGCGA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   281 AACCCGGCCTCACGGAGGCACTTACCGGGAATGTTATGGACACAGTTCAATCCTGGGGCTTGCGA 340

Qy   301 AATTTAATAAACCTGGGGAAAAATATTATGAGAAAGCTGTAAAACGCCTTATGATCCTTGCCAGGA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   341 AATTTAATAAACCTGGGGAAAAATATTATGAGAAAGCTGTAAAACGCCTTATGATCCTTGCCAGGA 400

Qy   361 AAAGCCTACTACGGATGGAGTCGGCCAAAGATCGGGTGAGATTGCCACTGGGTCCCCGTCA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   401 AAAGCCTACTACGGATGGAGTCGGCCAAAGATCGGGTGAGATTGCCACTGGGTCCCCGTCA 460

Qy   421 ACTGAACTGGGACAATGTCCCATAGAGATTTCAAGTACCCACACAAGAAACTCA 472
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   461 ACTGAACTGGGACAATGTCCCATAGAGATTTCAAGTACCCACACAAGAAACTCA 512

Search completed: March 20, 2006, 14:08:39
Job time : 1708 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:39:54 ; Search time 854 seconds
                                              (without alignments)
                                              16357.372 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctccgaagaa.........aaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332345308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :   N_Geneseq_21:*
                    1: geneseqn1980s:*
                    2: geneseqn1990s:*
                    3: geneseqn2000s:*
                    4: geneseqn2001as:*
                    5: geneseqn2001bs:*
                    6: geneseqn2002as:*
                    7: geneseqn2002bs:*
                    8: geneseqn2003as:*
                    9: geneseqn2003bs:*
                    10: geneseqn2003cs:*
                    11: geneseqn2003ds:*
                    12: geneseqn2004as:*
                    13: geneseqn2004bs:*
                    14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	2028.4	96.8	3617	13	ACN38684	Acn38684 Tumour-as
2	1821.4	86.9	2317	4	AAF32674	Aaf32674 Human as
3	1514.6	72.3	2471	12	ADM86946	Adm86946 Human pro
4	1448.4	69.1	1510	12	ADM87430	Adm87430 Human EST
5	1216.4	58.0	2120	13	ADR25944	Adr25944 Breast ca
6	1206.4	57.6	1221	14	ADV43099	Adv43099 Human psy
7	1071	51.1	1729	5	AAS73957	Aas73957 DNA encod
8	1021.2	48.7	1993	5	AAS73958	Aas73958 DNA encod
9	594	28.3	625	11	ADT94962	Adt94962 Colon can
10	594	28.3	625	11	ADX41444	Adx41444 Human cDN
11	576.4	27.5	1242	11	AAC77875	Aac77875 Human can
12	576.2	27.5	641	11	ADT95471	Adt95471 Colon can
13	576.2	27.5	641	11	ADX41953	Adx41953 Human cDN
14	556	26.5	559	6	ABK29726	Abk29726 Colon ade
15	520.8	24.8	539	11	ADT94802	Adt94802 Colon can
16	520.8	24.8	539	11	ADX41284	Adx41284 Human cDN
17	488	23.3	511	6	ABK70212	Abk70212 Human lun
18	470.4	22.4	512	6	ABK70219	Abk70219 Human lun
19	355	16.9	3128	12	ADG64787	Adg64787 Novel hum
20	329	15.7	449	5	AAS73955	Aas73955 DNA encod
21	274.8	13.1	439	5	AAS73956	Aas73956 DNA encod
22	259.6	12.4	2080	5	AAV72220	Aav72220 Human IRS
23	259.6	12.4	2080	10	ADD35681	Add35681 Human ins
24	259.6	12.4	2080	10	ADE86152	Ade86152 Human IRS
25	258	12.3	2129	11	ADR66181	Adr66181 Human pro
26	258	12.3	2129	13	ADX07028	Adx07028 Cyclin-de
27	258	12.3	2137	2	AAZ52899	Aaz52899 Human pro
28	258	12.3	2877	14	ADX07032	Adx07032 Cyclin-de
29	258	12.3	3168	14	ADX07030	Adx07030 Cyclin-de
30	258	12.3	3255	11	ADL22621	Adl22621 Human dis
31	258	12.3	3458	8	ABX34756	Abx34756 Human mdd
32	255.4	12.2	295	10	ABL79691	Abl79691 Human ova
33	241.6	11.5	2328	10	ADD35683	Add35683 Human ins
34	241.6	11.5	2328	10	ADE86154	Ade86154 Human IRS
35	240.2	11.5	45698	12	ADJ12490	Adj12490 DNA fragm
36	227.4	10.8	1415	14	ADZ77086	Adz77086 IRSp53 ri
37	219.6	10.5	2133	4	AAH18159	Aah18159 Human cDN
38	216.4	10.3	16181	4	ABK42539	Abk42539 Genomic s
39	216.4	10.3	16181	4	AAL04268	Aal04268 Human rep
40	216.4	10.3	16181	6	ADB60695	Adb60695 Connectiv
41	209.6	10.0	476	9	ABL67044	Abl67044 Thyroid c
42	209.6	10.0	476	9	ABL68013	Abl68013 Ovary can
43	209.6	10.0	476	9	ABL67457	Abl67457 Thyroid c
44	203.8	9.7	363	10	ADF66284	Adf66284 Human mic
45	203.8	9.7	363	13	ACF87869	Acf87869 Human SIR

ALIGNMENTS

RESULT 1
ACN38684
ID   ACN38684 standard; cDNA; 3617 BP.
XX
AC   ACN38684;
XX
DT   18-NOV-2004 (first entry)
XX
DE   Tumour-associated antigenic target (TAT) cDNA DNA324988, SEQ ID NO:2351.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Wu TD, Zhang Z, Zhou Y;
XX
DR   WPI; 2004-347921/32.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 1; SEQ ID NO 2351; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)

```
CC  polypeptides, and their related nucleic acids. The TAT polypeptides are
CC  overexpressed in cancer tissues compared to normal tissues, and may thus
CC  serve as effective targets for the diagnosis and treatment of cancer in
CC  mammals. The invention also relates to nucleic acid and polypeptide
CC  sequences at least 80% identical to the TAT nucleic acids and
CC  polypeptides; expression vectors and host cells comprising a TAT nucleic
CC  acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC  molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC  TAT polypeptide; and methods and compositions for the treatment or
CC  diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC  antibodies, antagonists, binding molecules and compositions are useful
CC  for diagnosing or treating a cell proliferative disorder associated with
CC  increased TAT expression, particularly cancers such as breast cancer,
CC  colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC  cancer, pancreatic cancer, cervical cancer, cancers of the central
CC  nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC  used as hybridisation probes, in chromosome and gene mapping, in
CC  chromosome identification and in gene therapy. The present sequence
CC  represents a TAT nucleic acid of the invention
XX
SQ  Sequence 3617 BP; 941 A; 941 C; 924 G; 811 T; 0 U; 0 Other;

Query Match             96.8%;  Score 2028.4;  DB 13;  Length 3617;
Best Local Similarity   99.9%;
Matches 2029;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy       1 CTCCTCTTGCTTCCTCGAAGAAAGGCCAGGGGGGCGGGGGCGCCCCGCAAGTTTTGACATTTTCGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTCCTCTTGCTTCCTCGAAGAAAGGCCAGGGGGGCGGGGGCGCCCCGCAAGTTTTGACATTTTCGCAG 60

Qy      61 CGGAGACGCGCGGGGGGCGCACTCTCTCGGGCGCCCGCGGGGGCGCGGGGCCCGACCCTCCAGAGC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGGAGACGCGCGGGGGGCGCACTCTCTCGGGCGCCCGCGGGGGCGCGGGGCCCGACCCTCCAGAGC 120

Qy     121 CCCTTAGTTCGTCGGCGGCCCCCTCCCCGGAGTCCCGGGCGGCCGCCGCCGCCAGCCGCCCAGCCG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CCCTTAGTTCGTCGGCGGCCCCCTCCCCGCTCCCGGGCGGCCATGTCCGGGCGGCCACCGCCCAGCCG 180
```

(sequence alignment continues)

```
Qy   1801 ATGATGGGGGGGCCATCCTGTCTGCGGCTGCCACCGTGCTGCGGGAGCTTCAGTGGGAGGAGGCCT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATGATGGGGGGGCCATCCTGTCTGCGGCTGCCACCGTGCTGACGGTGGGAGGAGGCCT 1860

Qy   1861 AACTCTAAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTCGTTTCACGTAGTTGGGTTGAC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AACTCTAAATGTCGCCTGCTTAAGCAAATCATGCTTCTCTCGTTTCACGTAGTTGGGTTGAC 1920

Qy   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCCATTTAAAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCCATTTAAAT 1980

Qy   1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAAACTTGGTTTCAAATCTTTA 2030
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATTTTCTTCCTATTCTGTTCAAGAAACAGTAAAACTTGGTTTCAAATCTTTA 2030
```

RESULT 2
AAF32674

```
ID   AAF32674 standard; cDNA; 2317 BP.
XX
AC   AAF32674;
XX
DT   22-MAR-2001  (first entry)
XX
DE   Human cDNA encoding intracellular signalling molecule INTRA37.
XX
KW   Human; intracellular signalling molecule; INTRA; immunosuppressive;
KW   cytostatic; neuroprotective; nootropic; antiarteriosclerotic; cancer;
KW   antiinflammatory; anti-HIV; neuroleptic; antibacterial; antifungal;
KW   antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;
KW   cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;
KW   inflammatory disorder; Addison's disease; gastrointestinal disorder;
KW   neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
KW   mental disorder; schizophrenia; anxiety; ss.
XX
OS   Homo sapiens.
XX
PN   WO20007040-A2.
XX
PD   21-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US016636.
XX
PR   16-JUN-1999;  99US-0139566P.
PR   17-AUG-1999;  99US-0149640P.
PR   09-NOV-1999;  99US-0164417P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Hillman JL,  Lal P,  Bandman O,  Baughn MR;
PI   Azimzai Y,  Yang J,  Reddy R,  Lu DAM;
XX
DR   WPI; 2001-025334/03.
DR   P-PSDB; AAB64405.
XX
PT   New human intracellular signalling molecules, useful for the diagnosis,
PT   prevention and treatment of cell proliferative, autoimmune, inflammatory,
PT   neurological, gastrointestinal, reproductive and developmental disorders.
XX
PS   Claim 5; Page 182-183; 192pp; English.
XX
CC   Sequences AAF32638 - AAF32689 represent cDNA encoding human intracellular
CC   signalling molecules INTRA1 - INTRA52, represented in AAB64369 -
CC   AAB64420. Modulators of the intracellular signalling molecules of the
CC   invention exhibit immunosuppressive; cytostatic; neuroprotective;
CC   nootropic; antiarteriosclerotic; antiinflammatory; anti-HIV; neuroleptic;
CC   antibacterial; antifungal; antiviral; antiparasitic; antihelminthic; and
CC   antiparkinsonian activity. INTRA polypeptides their agonists and
CC   antagonists are useful for the treatment of a condition associated with
CC   decreased or increased expression of functional INTRA. Disorders
CC   associated with abnormal INTRA expression or activity include cell
CC   proliferative disorders e.g. arteriosclerosis and cancers; autoimmune or
CC   inflammatory disorders e.g. Addison's disease and acquired
CC   immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections; gastrointestinal disorders e.g.
CC   dysphagia and irritable bowel syndrome; neurological disorders e.g.
CC   epilepsy and Parkinson's disease; prion diseases e.g. Creutzfeldt-Jakob
CC   disease and mental disorders e.g. anxiety, schizophrenia and Tourette's
CC   disorder. Antibodies immuno specific for the INTRA proteins may also be
CC   useful in the diagnosis of the above disorders
XX
SQ   Sequence 2317 BP; 676 A; 554 C; 545 G; 542 T; 0 U; 0 Other;
```

```
Query Match     86.9%;  Score 1821.4;  DB 4;  Length 2317;
Best Local Similarity   99.9%;  Pred. No. 2.3e-304;
Matches 1822;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
```

```
Qy    208 GCCGCAGCCATGTCCCGGGCGCCCGAGGAGGTGAACCGGCTCACGGAGGCACCTACCGG 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     30 GCCGCAGCCATGTCCCGGGCGCCCGAGGAGGTGAACCGGCTCACGGAGGCACCTACCGG 89

Qy    268 AATGTTATGGAACAGTTCAATCCTCGGGGCTGCCGAAATTTAATAAAACCTGGGGAAAATTAT 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     90 AATGTTATGGAACAGTTCAATCCTGGGGCTGCCGAAATTTAATAAAACCTGGGGAAAAATTAT 149

Qy    328 GAGAAAGCTGTAAACCCTATATCCTTCTTGCGAGAAAAGCCTACTACCATGGACTGGCCAAG 387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    150 GAGAAAGCTGTAAACCCTATATCCTTGCGAGAAAAGCCTACTACCATGGACTGGCCAAG 209

Qy    388 ATCGGTGAGATTGGCCACTGGGTTCCCCCGTGTCAACTGAACTGGGACATGTCTCATCAGAG 447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 ATCGGTGAGATTGGCCACTGGGTTCCCCGTGTCAACTGAACTGGGACATGTCTCATCAGAG 269

Qy    448 ATTTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAAATTCCAC 507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 ATTTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGATGAAAATTTTAAAAAAATTCCAC 329

Qy    508 AAAGAGATTATCCATGAGCTGGAAGAAGAAGATGAACTTGACGTGAAATATATGAACGCA 567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 AAAGAGATTATCCATGAGCTGGAAGAAGAAGATGAACTTGACGTGAAATATATGAACGCA 389

Qy    568 ACTCTAAAAAGATACCAAAACAGAACAAGAATAAATTAGAGTCTTTGGAGAAAATCCAA 627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 ACTCTAAAAAGATACCAAAACAGAACAAGAATAAATTAGAGTCTTTGGAGAAAATCCAA 449

Qy    628 GCTGAGTTGAAGAAGATCAGAAGGAAAAAGCCAAGAAACCGACACTCAAATATGAA 687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 GCTGAGTTGAAGAAGATCAGAAGGAAAAAGCCAAGAAACCGACACTCAAATATGAA 509

Qy    688 CACAAAGAAAATTGAGTATGTGGAGACCGTTACTTCGTCAGAGTGAAAATCCAGAAATTC 747
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 CACAAAGAAAATTGAGTATGTGGAGACCGTTACTTCGTCAGAGTGAAAATCCAGAAATTC 569

Qy    748 ATTTGCAGATGGTTGCAAAGAGGGCTTCTTGCTTGAAGACGAAGAAGGCGCTTTCTGGTT 807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 ATTTGCAGATGGTTGCAAAGAGGGCTTCTTGCTTGAAGACGAAGAAGGCGCTTTCTGGTT 629

Qy    808 GATAAGCACTGTGGCCTTTGCAAACACAACTACATTATATACTTACAGTCTGCCAGACTA 867
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 GATAAGCACTGTGGGCTTTGCAAACACCATACATTATTATCACTTACAGTCTGCAGACTA 689

Qy    868 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCGTGTTGTTGATGCCATCAAAGTGCCAGAG 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 CTGAATTCCAAGCTGCCTCGGTGGCAGGAGACCGTGTTGTTGATGCCATCAAAGTGCCAGAG 749

Qy    928 AAAATCATGAATATGATCTGGAAGAAAATAAAAGACCCCAGCCTCTACCCCGTGTTCTGGAACT 987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 AAAATCATGAATATGATCGGAAGAAATAAAAGACCCCAGCCTCTACCCCGTGTTCTGGAACT 809

Qy    988 CCTCAGGCTTCACCCATGATCGAGAGGAGAGCAATGTGGTTAGGAAGAAGATTACGAACACCCTT 1047
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 CCTCAGGCTTCACCCATGATCGAGAGGAGAAGCAATGTGGTTAGGAAGAAGATTACGAACACCCTT 869

Qy   1048 TCTAAAATGCTCACCAAAGAATGCCCCCCGCTCCTTCAGGCAGGCAGAGCATATACCAGTCCCTTG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 3
ADM86946
ID   ADM86946 standard; cDNA; 2471 BP.
XX
AC   ADM86946;
XX
XX   ADM86946;
```

```
Db    870 TCTAAATGCTCACCAAAGAATGACCCCCGGCGTTCCTTCAGGCGAGAGCGATATACCAGTCCCTTG 929
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1108 ATCGGATATGTTTAATAAACCCAGCCACGGCTGCCCCCGGAATTCACAAAGGGTAAATAATTCA 1167

Db    930 ATCGGATATGTTTAATAAACCCAGCCACGGCTCCCCCCGAATTCACAAAGGGTAAATAATTCA 989
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1168 ACAGGTACTTCCGAAGATCCCAGTTTACACGGATCCAGTTTCGGTTGCAACGGGGACTGAAC 1227

Db    990 ACAGGTACTTCCGAAGATCCCAGTTTACACGGATCAGTTTCGGTTGCAACGGGGACTGAAC 1049
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1228 ATGATGAAGAAGAAGCAGAAAGTGAAGAGAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1287

Db   1050 ATGATGAAGAAGAAGCAGAAAGTGAAGAAACCATCTTCCCGCACACTGCGGGCTCCAACAAGACC 1109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1288 TTACTCAGCTTTGCACAGGGAGGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1347

Db   1110 TTACTCAGCTTTGCACAGGGAGGATGTCATCACGCTGCTCATCCCCGAGGAGAAGGATGGC 1169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1348 TGGCTCTATGGGAGGACACGACGTGTCCAAGGCGGGGGTTGGTTCCGGTTGCGTTACACG 1407

Db   1170 TGGCTCTATGGGAGGACACGACGTGTCCAAGGCGGGGGTTGGTTCCGGTTGCGTTACACG 1229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1408 AAGTTGCTGGAAGAAAATGAGACAGAAGAAGAGTGACCGTGCCACCGCCAAGCCCACACCA 1467

Db   1230 AAGTTGCTGGAAGAAAATGAGACAGAAGAAGAGTGACCGTGCCACCGCCAAGCCCACACCA 1289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1468 GTGAGAAGCATCAGCACCGTGAAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1527

Db   1290 GTGAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCC 1349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1528 GACTACTTGGAAATGCTTGTTCCATGGGGGACAGCTGCGACAGGAGAGAGCAGAATTCGGCCAGG 1587

Db   1350 GACTACTTGGAAATGCTTATCCATGGGGGCAGCTGCGACAGGAGAGCAGAATTCGGCCAGG 1409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1588 ACGACAATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCCGAGGCTCCTAACGATGCC 1647

Db   1410 ACGACAATCCACCTTTAAGGCCCCAGCGTCCAAGCCCGAGACCCGAGGCTCCTAACGATGCC 1469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1648 AACGGGACTGCAAAGCCGCCCTTTGCACTGTGAAACTC 1707

Db   1470 AACGGGACTGCAAAGCCGCCTTTTCTCAGGGAGAAAACCCCTTTGCCACTGTGAAACTC 1529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1708 CGCCCGACTGTGACGAATGATCGGCTCGGCACCCATCATTCGATGAGAGCCAAGGA 1767

Db   1530 CGCCCGACTGTGACGAATGATCGGCTCGGCACCCATCATTCGATGAGAGCCAAGGA 1589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1768 CTCTCCCCGGGCCTCTCCCGGTTCATCTCGGGAATGATGATGGGGCGCATCCTGTCTCTGCCACGT 1827

Db   1590 CTCTCCCGGGCCTCTCCGGTTCTCCGCAAATGATGGGGCCATCCTGTCTCTGCCACGT 1649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1828 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTTAAGCAAA 1887

Db   1650 GCTGACGGTCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAATGTCGCCTTAAGATAAATGAGT 1709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1888 TCATGCTTCTCTGTTTCACGTAGTTGCAAGTTTCTGCGTCTTTAAGATAAATGAGT 1947

Db   1710 TCATCGCTTCTCTGTTTCACGTAGTTGCAAGTTTCTGACAAGTTCTGCCCTTAAGATAAAATGAGT 1769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1948 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTTCTTCTTAGTTCTGTTCAAGAAAC 2007

Db   1770 AATAGTCTAATGACCAGCTCAGCCATTTAAAATATTTTTCTTCTTAGTTCTGTTCAAGAAAC 1829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2008 AGTAAACTTGGTTTCAATCTTTA 2030

Db   1830 AGTAAACTTGGTTTCAATCTTTA 1852
          |||||||||||||||||||||||
Qy   ...  AGTAAACTTGGTTTCAATCTTTA 1852
```

```
CC   human polynucleotide sequence from the present invention. N.B. The
CC   sequences for this patent were obtained from the USPTO web site from an
CC   equivalent US patent US20040048249A1.
XX
SQ   Sequence 2471 BP; 681 A; 634 C; 585 G; 571 T; 0 U; 0 Other;

     Query Match     72.3%;  Score 1514.6;  DB 12;  Length 2471;
     Best Local Similarity   99.7%;  Pred. No.1.6e-251;
     Matches 1528;  Conservative   0;  Mismatches   4;  Indels   1;  Gaps   1;

Qy    499 AAATTCCACAAAGAGGATTATTCCATGAGCTCGGAGGAAGAGATAGAACTTGACGTGAAATAT  558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAATTCCACAAAGAGGATTATTCCATGAGCTCGGAGGAAGAGATAGAACTTGACGTGAAATAT   60

Qy    559 ATGAACGCAACTCTAAAAAAGATACCAAACAGAACAGAAGAATAAATTAGAGTTCTTTGGGAG  618
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGAACGCAACTCTAAAAAAGATACCAAACAGAACAGAAGAATAAATTAGAGTTCTTTGGAG  120

Qy    619 AAATCCCAAGCTGAGTTGAAGAAGATCAGAGGAGAAAAGCCAAGACGCAAACGCACTC  678
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAATCCCAAGCTGAGTTGAAGAAGATCAGAGGAGAAAAGCCAAGACCCAAACGCACTC  180

Qy    679 AAATATGAACACAAAAGAAATTGAGTTGTGGGAGACCGTTACTTCTTCGTCAGAGTGAAATC  738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAATATGAACACAAAAGAAATTGAGTTGTGGGAGACCGTTACTTCTTCGTCAGAGTGAAATC  240

Qy    739 CAGAAATTCATTGCAGATAGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGACGCGTTCTGC  798
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGAAATTCATTGCAGATAGGTTGCAAAGAGGCTCTGCTTGAAGAGAAGACGCGGCTTCTGC  300

Qy    799 TTTCTGGTTGATAAGCACTGTGGCCTTTGCAAACCACATACATTAATTATATCACTTACAGTCT  858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTCTGGTTGATAAGCACTGTGGCCTTTGCAAACCACATACATTAATTATATCACTTACAGTCT  360

Qy    859 GCAGAACTACTGAATTCCAAAGCTGCCTCGGTGGGGCAGGAGGCACCTGTTGTTGCCATCAAA  918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCAGAACTACTGAATTCCAAAGCTGCCTCGGTGGGGCAGGAGGCACCTGTTGTTGCCATCAAA  420

Qy    919 GTGCCCAGAGAAAATCATGAATATGATCGAAGAGAATAAAGACCCAGCCTTCTACCCCCGTG  978
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGCCCAGAGAAAATCATGAATATGATCGAAGAGAATAAAGACCCAGCCTTCTACCCCCGTG  480

Qy    979 TCTGGAACTCCTCAGGGCTTCACCCATGATCGAGAGAGAAGAGCAATGTGGTTAGGAAGAGTTAC 1038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCTGGAACTCCTCAGGGCTTCACCCATGATCGAGAGAGAAGAGCAATGTGGTTAGGAAGAGTTAC  540

Qy   1039 GACACCCTTTCTAAATGCTCACCAAAAGATGCCCCCCGCTTCCTTCAGGCGAGAGCATATACC 1098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACACCCTTTCTAAATGCTCACCAAAAGATGCCCCCCGCTTCCTTCAGGCGAGAGCATATACC  600

Qy   1099 AGTCCCTTGATCGATATGTTTAATACCAGCCACGGCCTGCCCCGAATTCACAAAGGGTA 1158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGTCCCTTGATCGATATGTTTAATACCAGCCACGGCCTGCCCCGAATTCACAAAGGGTA  660

Qy   1159 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACACGGATCAGTTGGTTGCAACG 1218
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACACGGATCAGTTGGTTGCAACG  720

Qy   1219 GGACTGAACATGATGAAGAGAGCAGAAAGTGAACGAACCATCATCTTCCCGCACACTGCGGGGCTCC 1278
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGACTGAACATGATGAAGAGAGCAGAAAGTGAACGAACCATCTTCCCGCACACTGCGGGGCTCC  780

Qy   1279 AACAAGACCTTACTCACGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCGAGGAG 1338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AACAAGACCTTACTCACGCTTTGCACAGGGAGATGTCATCACGCTGCTCATCCCGAGGAG  840

Qy   1339 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGACGAGGGTGGTTGGTTCCCGTCG 1398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAGGATGGCTGGCTCTATGGAGAACACGACGTGTCCAAGGACGAGGGTGGTTGGTTCCCGTCG  900

Qy   1399 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCCAAGC 1458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCGTACACGAAGTTGCTGGAAGAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCCAAGC  960

Qy   1459 CCCACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGCCGGCAGCTGCCCGCGCTCTC 1518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGCCGGCAGCTGCCCGCGCTCTC 1020

Qy   1519 CCCCACCCCGACTACTTGGAATGCTTATCCAATGGGGCCCCAGAGACCCGGCGCTCCT 1578
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCACCCCGACTACTTGGAATGCTTATCCAATGGGGCCCCAGAGACCCGGCGCTCCT 1080

Qy   1579 TCGGTCAGGAGGACGACATCCACCTTTAAGGCCCCAGGGCCGGTCTCCAAAGCCAAGCCGCCGGGGCTCCACT 1638
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCGGTCAGGAGGACGACATCCACCTTTAAGGCCCCAGTGTTCAAGGCCGAGAAAACCCCTTTGCCACT 1140

Qy   1639 AACGGATGGCCAAGGGACTGCAAAGCCGCCTTTTCTCAGGCGAAGAAAACCCCTTTGCCACT 1698
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AACGGATGGCCAAGGGACTGCAAAGCCGCCTTTTCTCAGGCGAAGAAAACCCCTTTGCCACT 1200

Qy   1699 GTGAAACTCCGCCCGACTGTGACGAAGAGTTCAGTGGAGAACGCTGGGCACCCATCATTCGATTGAGGAGGAC 1758
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGAAACTCCGCCCGACTGTGACGAAGAGTTCAGTGGAGAACGCTGGGCACCCATCATTCGATTGAGGAGGAC 1260

Qy   1759 AGCCAAGGACTCTCCCGGGGGCCTCTCCCGGGTTCCCTTGCGGAATGAGTGGGGGGCCATCCTGT 1818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGCCAAGGACTCTCCCGGGGGCCTCTCCCGGGTTCCCTTGCGGAATGAGTGGGGGGCCATCCTGT 1320

Qy   1819 CTGCCACAGTGCTGCGACGGTCGGGAGAAGCTTCAGTGGAGGAGGGGCCTTAACTCTAATGTTCGCGCTGC 1878
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGCCACAGTGCTGCGACGGTCGGGAGAAGCTTCAGTGGAGGAGGGGCCTTAACTCTAATGTTCGCGCTGC 1380

Qy   1879 TTAAGGAAAATCATG-CTTCCTCTGTTTCACGTAGTTGGGTTGGGTTGACCAAGTTTCTGCCCTTTAAG 1937
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TTAAGGAAAATCATGGCTTCCTCTGTTTCACGTAGTTGGGTTGGGTTGACCAAGTTTCTGCCCTTTAAG 1440

Qy   1938 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCCATTTAAAATAATATTTTCTTCCTATTCTG 1997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATAAATGAGTAATAGTCTAATGACCAGCTCAGCCCATTTAAAATAATATTTTCTTCCTATTCTG 1500

Qy   1998 TTCAAGAAACAGTAAAACTTGGTTTCAATCTTTA 2030
          ||||||||||||||||||||||||||||||||||
Db   1501 TTCAAGAAACAGTAAAACTTGGTTTCAATCTTTA 1533
```

RESULT 4
ADM87430

```
ID   ADM87430 standard; cDNA; 1510 BP.
XX
AC   ADM87430;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human EST derived nucleotide sequence SEQ ID NO:523.
XX
KW   respiratory; cytostatic; antiarthritic; antiinflammatory;
KW   gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW   antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW   chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW   inflammatory condition; arthritis; inflammatory bowel disease;
KW   Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW   graft versus host disease; human; expressed sequence tag; EST; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200409834-A2.
XX
PD   29-JAN-2004.
XX
PF   19-JUL-2002; 2002WO-US022858.
XX
PR   21-JUL-2001; 2001US-0306971P.
PR   28-MAR-2002; 2002US-00112944.
XX
PA   (NUVE-) NUVELO INC.
XX
```

```
PI   Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
PI   Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
XX
DR   WPI; 2004-143291/14.
DR   P-PSDB; ADM87648.
XX
PT   New isolated polynucleotides and polypeptides, useful for treating, e.g.
PT   cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
PT   Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
PT   versus host disease.
XX
PS   Example 2; SEQ ID NO 523; 591pp; English.
XX
CC   The present invention describes an isolated polynucleotide (I): (a)
CC   comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
CC   which encodes a polypeptide with biological activity, where the
CC   polynucleotide hybridises to (i) under stringent hybridisation conditions
CC   or has greater than 99% sequence identity with (i). (I) has respiratory,
CC   cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
CC   antibacterial, immunosuppressive, antidiabetic and antirheumatic
CC   activities, and can be used in gene therapy. (I) can be used for
CC   generating polynucleotides encoding chimeric or fusion proteins and
CC   heterologous protein sequences. The polynucleotides can be used to
CC   express recombinant protein for analysis, characterisation or therapeutic
CC   use; as markers for tissues in which the corresponding protein is
CC   preferentially expressed; as molecular weight markers on gels; as
CC   chromosome markers or tags to identify chromosomes or to map related gene
CC   positions; to compare with endogenous DNA sequences as probes to
CC   identify potential genetic disorders; as probes to hybridise and discover
CC   genes, related DNA sequences; as a source of information to derive PCR
CC   primers for genetic fingerprinting; as a probe to subtract-out known
CC   sequences in the process of discovering other novel polynucleotides; for
CC   selecting and making oligomers for attachment to a gene chip or other
CC   support, including for examination of expression patterns; to raise anti-
CC   protein antibodies using DNA immunisation techniques; and as an antigen
CC   to raise anti-DNA antibodies or elicit another immune response. The
CC   polynucleotides and polypeptides can also be used as nutritional sources
CC   or supplements, e.g. as a protein or amino acid supplement, as a carbon
CC   source, as a nitrogen source or as a source of carbohydrates. The
CC   polynucleotides and polypeptides can also be used treat cancer. The
CC   compositions are useful for promoting better or faster closure of non-
CC   healing wounds, for the generation and regeneration of tissues, for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, and conditions resulting from
CC   systemic cytokine damage. The compositions can also be used to treat
CC   inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
CC   Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
CC   or graft versus host disease. The present sequence represents an
CC   expressed sequence tag (EST) derived nucleotide sequence from the present
CC   invention. N.B. The sequences for this patent were obtained from the
CC   USPTO web site from an equivalent US patent US20040482491.
XX
SQ   Sequence 1510 BP; 462 A; 372 C; 366 G; 310 T; 0 U; 0 Other;.

Query Match             69.1%;  Score 1448.4;  DB 12;  Length 1510;
Best Local Similarity   99.2%;
Matches 1498;  Conservative    0;  Mismatches    6;  Indels    6;  Gaps   4;

Qy    217 ATGTCCCGGGGGGCCCCGAGGAGGTGAACCGCGCTCACGGAGAGCACCTACCGGAATGTTATG 276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTCCCGGGGGGCCCCGAGGAGGTGAACCGCGCTCACGGAGAGCACCTACCGGAATGTTATG 60

Qy    277 GAACAGTTCAATCCTCGGGCTCGGAAAATTTAATAAACCTGGGAAAAATTATGAGAAAGCT 336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAACAGTTCAATCCTGGGCTCGGAAAATTTAATAAACCTGGGAAAAATTATGAGAAAGCT 120

Qy    337 GTAAACGCTATGATCCTGCGGCAGGAAAAAGCCTACTACGATGGAATCGGTGAG 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAAACGCTATGATCCTGCGGCAGGAAAAAGCCTACTACGATGGAATCGGTGAG 180

Qy    397 ATTGCCACTGGGTCCCCGTGTCAAC-TGAACTGGGGACCATGTCCTCATAGAGATTCAAG 455
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    181 ATTGCCACTGGGTCCCCGTGTCAACTGGGGACCGTTCCTCATAGAGATTCAAG 240

Qy    456 TACCCACAAGAAAAC-TCAACGAGGAGTCTT---GATGAAAATTTTAAAAAAATTCCACAAAG 511
          ||||||||||||||| ||||||||||||||    ||||||||||||||||||||||||||||
Db    241 TACCCACAAGAAAACTTCAACGAGGAGTCTTTGATGAAAATTTTAAAAAAATTCCACAAAG 300

Qy    512 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTT-GACGTGAAATATATGAACGCAACT 570
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    301 AGATTATCCATGAGCTGGAGAAGAAGATAGAACTTTGACGTGAAATATATGAACGCAACT 360

Qy    571 CTAAAAAGATACCAAACAGAACACAGAAGAATAAATTAGAGTCTTTGGAGAAATCCAAGCT 630
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAAAAAGATACCAAACAGAACACAGAAGAATAAATTAGAGTCTTTGGAGAAATCCAAGCT 420

Qy    631 GAGTTGAAGAAGATCAGAAGAGGAGAAAAGCCAAGGAAGAGGCCACTCAAAATATGAACAC 690
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAGTTGAAGAAGATCAGAAGAGGAGAAAAGCCAAGGAAGAGGCCACTCAAAATATGAACAC 480

Qy    691 AAAGAAATTGATGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 750
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAAGAAATTGATGTATGTGGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAATTCATT 540

Qy    751 GCAGAATGGTTGCAAAGAGCGCTTCGTGCTTGAAGAGAAGAGAGGCGCTTCTGCTTTCTCGGTTGAT 810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCAGAATGGTTGCAAAGAGCGCTTCGTGCTTGAAGAGAAGAGAGGCGCTTCTGCTTTCTCGGTTGAT 600

Qy    811 AAGCACTGTGGCTTTGCAAACCACATACATTATTATTCACTTACAGTCTGCAGAACTACTG 870
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGCACTGTGGCTTTGCAAACCACATACATTATTATTCACTTACAGTCTGCAGAACTACTG 660

Qy    871 AATTCCAAGCTGCCTCGGTGTGGCAGGAGACCGTGTTGTTGATGCCATCAAAGTGCCAGAA 930
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATTCCAAGCTGCCTCGGTGTGGCAGGAGACCGTGTTGTTGATGCCATCAAAGTGCCAGAA 720

Qy    931 ATCATGAATATGATCGAAGAAAATAAAGACCCCAGCCTCTACCCCCGTGTGTTGGGAACTCCT 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ATCATGAATATGATCGAAGAAAATAAAGACCCCAGCCTCTACCCCCGTGTGTTGGGAACTCCT 780

Qy    991 CAGGCCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAGAAGATTACACACCCTTTCT 1050
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGGCCTTCACCCATGATCGAGAGAAGCAATGTGGTTAGGAGAAGATTACACACCCTTTCT 840

Qy   1051 AAATGCTCACCAAAGATGCCCCCGCTGCCCGAATTCACAAAGGGTAAAATAATTCAACA 1110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAATGCTCACCAAAGATGCCCCCGCTGCCCGAATTCACAAAGGGTAAAATAATTCAACA 900

Qy   1111 GATATGTTTAATAACCCAGCCACGGCTGCCCGGAATTCACAAAGGGTAAAATAATTCAACA 1170
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GATATGTTTAATAACCCAGCCACGGCTGCCCGGAATTCACAAAGGGTAAAATAATTCAACA 960

Qy   1171 GGTACTTCCCGAAGATCCCAGTTTACAGCGACATCAGTTTCGGTTGCAAACGGGACTGAACATG 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGTACTTCCCGAAGATCCCAGTTTACAGCGACATCAGTTTCGGTTGCAAACGGGACTGAACATG 1020

Qy   1231 ATGAAGAAGACAGAAAGTGAAGACCAATCTTCCCGCACACTGCGGGGCTTCCAACAAGACCTTA 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATGAAGAAGACAGAAAGTGAAGACCAATCTTCCCGCACACTGCGGGGCTTCCAACAAGACCTTA 1080

Qy   1291 CTCAGCTTTGCACAGGGAGAGTCATCACGCTGCTCATCCCGGAGGAGAAGGATGGCTGG 1350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTCAGCTTTGCACAGGGAGAGTCATCACGCTGCTCATCCCGGAGGAGAAGGATGGCTGG 1140

Qy   1351 CTCTATGGGAGAACACGACGTGTCCAAGGGCAGCGAGGGGTTGGTTCCCGTGCTGTTGCACGGAAG 1410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTCTATGGGAGAACACGACGTGTCCAAGGGCAGCGAGGGGTTGGTTCCCGTGCTGTTGCACGGAAG 1200

Qy   1411 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACGTGCCCAGCGCCGTGACCACACCACCAGTG 1470
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TTGCTGGAAGAAAATGAGACAGAAGCAGTGACGTGCCCAGCGCCGTGACCACACCACCAGTG 1260

Qy   1471 AGAAGCAATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTTGCATGTCACCCCCGAC 1530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGAAGCAATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTTGCATGTCACCCCCGAC 1320
```

```
Qy   1531 TACTTGGAATGCTTGTCCATGGGGGGGGGGCACCAGGAGGAGCAGGATTCGGCCAGGACG 1590
          ||||||||| |||||| ||||||||||||| ||| || ||||||||||||| ||||||||
Db   1321 TACTTGGAAATGCTTATCCATGGGGGGGCGCAGCTGCCGCGCGACAGGAGGAGATTCGGCCACTGGGACG 1380

Qy   1591 ACATCCACCTTTAAGGCCCCCAGGCGTCCAAGCCCCGAGACGCGGCCGCACAGGAGGACG 1650
          ||||||||||||||||||||||||| |||| |||||||||||||||||||||||||||||
Db   1381 ACATCCACCTTTAAGGCCCCCAGGCGTCCAAGCCCCGAGACGCGGCCGCACAGGAGGACG 1440

Qy   1651 GGGGACTGCAAAGCGGCCCTTTTCTCCAGCGCGAGAAAACCCCTTTGGGCGAGAAAACTC 1710
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GGGGACTGCAAAGCGCCCTTTTCTCCACTGTCGACTGTCCTCATAGAGATTCCAAGTACC 1500

Qy   1711 CCGACTGTGA 1720
          ||||||||||
Db   1501 CCGACTGTGA 1510
```

RESULT 5

```
ID   ADR25944 standard; DNA; 2120 BP.
AC   ADR25944;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Breast cancer prognosis marker #1805.
XX
KW   ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200406545-A2.
XX
PD   05-AUG-2004.
XX
PF   15-JAN-2004; 2004WO-US001100.
XX
PR   15-JAN-2003; 2003US-00342887.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
PA   (NECA-) NETHERLANDS CANCER INST.
XX
PI   Van't Veer LJ, He Y;
XX
DR   WPI; 2004-593473/57.
XX
PT   Classifying a breast cancer patient according to prognosis comprises
PT   determining the similarity between the level of expression of each of
PT   five genes in a cell sample taken from patient, to control levels.
XX
PS   Disclosure; SEQ ID NO 1805; 226pp; English.
XX
CC   The invention relates to a method of classifying a breast cancer patient
CC   according to prognosis by determining the similarity between the level of
CC   expression of each of five genes for which markers are listed in the
CC   specification, in a cell sample taken from the breast cancer patient, to
CC   control levels of expression for each respective five genes to obtain a
CC   patient similarity value. The methods are useful for classifying a breast
CC   cancer patient according to prognosis. Kits and computer program products
CC   are useful for data analysis using the diagnostic, prognostic and
CC   statistical methods of the invention. This sequence corresponds to a
CC   marker used in the method of the invention.
XX
SQ   Sequence 2120 BP; 625 A; 513 C; 464 G; 518 T; 0 U; 0 Other;
```

```
Query Match    58.0%;  Score 1216.4;  DB 13;  Length 2120;
Best Local Similarity   98.3%;  Pred. No. 3.7e-200;
Matches 1229;  Conservative   0;  Mismatches  21;  Indels   0;  Gaps   0;
```

```
Qy    292 GGGGCTGCGAAATTTAATAAACCTGGGGGAAAAATTATGAGAAAAGCTATGATC 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 GGGGCTGCGAAATTTAATAAACCTGGGGGAAAAATTATGAGAAAAGCTATGATC 411

Qy    352 CTGGCGAGGAAAAGCCTACTTACGATGGAGTGGCCAAGATCCGGTGAGATTGCCAGTGATC 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 CTGGGCAGGAAAAGCCTACTTACGATGGAGTGGCCAAGATCGGTGAGATTGCCACTGGGTCC 471

Qy    412 CCCGTGTCAACTGAACTGGGACATGTTCCTCATAGAGATTTCAAGTACCCCACAACAGTC 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 CCCGTGTCAACTGAACTGGGACATGTTCCTCATAGAGATTTCAAGTACCCACACAACTC 531

Qy    472 AACGGAGAGTCTTGATGAAAATTTTAAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    532 AACGGAGAGTCTTGATGAAAATTTTAAAAAAATTCCACAAAGAGATTATCCATGAGCTGGAG 591

Qy    532 AAGAAGATAGAAACTTGACGTGAAATATATGAACCGAACTTCTAAAAAGATACCAAACAGAA 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    592 AAGAAGATAGAACTTGACGTGAAATATATGAACCGAACTTCTAAAAAGATACCAAACAGAA 651

Qy    592 CACAAGAAGATAAAATTAGAGTCTCTTTGGGAGAAATCCCAAGCTGAGTTGAACAGAACG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 CACAAGAATAAAATTAGAGTCTTTTGGGAGAAATCCCAAGCTGAGTTGAACAGAACG 711

Qy    652 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATATGAACAGAAAGTGAGTTATGTGGAG 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    712 AAAAGCCAAGGAAGCCGAAACGCACTCAAATATGAACACAAAAGAAATTGAGTATGTGGAG 771

Qy    712 ACCGGTTACTTCTCGTCAGAGTGAAAATCCAGAAAATCATTGCAGAATGGGTTGCAAAGAGGCT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    772 ACCGGTTACTTCTCGTCAGAGTGAAAATCCAGAAATTCATTGCAGATGGGTTGCAAAAGCT 831

Qy    772 CTGCTTGAAGAGAAGAGCGCTCTCGCTTCTGTTGATAAGAGCACTGTGGCTTGCAAAC 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    832 CTGCTTGAAGAGAAGAGGCGCTTCGTTCTGCTTTCGGTTGATAAGGCACTGTGGCTTTGCAAAC 891

Qy    832 CACATACATTATTATTCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTGG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 CACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCCAAGCTGCCTCGGTCG 951

Qy    892 CAGGAGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATGATGATGCGAGAA 951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    952 CAGGAGAGACCTGTGTTGATGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATCGGAGAA 1011

Qy    952 ATAAAGACCCCAGCCCTTACCCCGTGTCTGGAAACTCCTCAGCGTTCACCCATGATCCAG 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1012 ATAAAGACCCCAGCCCTTACCCCGTGTCTGGAAACTCCTCAGCGCTTCACCCATGATCCAG 1071

Qy   1012 AGAAGCAATGTGGTTAGGAAAGATTTACGACACCCTTTCTAAAATGCTTCACCAAAGATGCCC 1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1072 AGAAGCAATGTGGTTAGGAAAGATTTACGACACCCTTTCTAAAATGCTTCACCAAAGATGCCC 1131

Qy   1072 CCCGGCTCCTTCAGGCAGCATATACCAGTGCCCTTGATCGATGATGTTTAATAACCAGCC 1131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1132 CCCGCTCCTTCAGGCAGCATATACCAGTGCCCTTGATCGATGATGTTTAATAACCCAGCC 1191

Qy   1132 ACGGCTGGCCCCGAATTCACAAAGGGTAAATATAAATTCAACAGGTGACTTTCCGGAAGATCCCAGT 1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1192 ACGGCTGGCCCCGAATTCACAAAGGGTAAATATAATTCAACAGGTACTTCCGGAAGATCCCAGT 1251

Qy   1192 TTACAGCGGCATCAGTTTCGGTTGCCAACGGGACTGAACATGATGAAGAGAAGAGTGAAG 1251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1252 TTACAGGGGATCAGTTTCGGTTGCCAACGGGACTGAACACATGATGAAGAGAAGAGTGAAG 1311

Qy   1252 ACCATCTTCCCGCACACTGCGCGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGGAGAT 1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1312 ACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGAT 1371

Qy   1312 GTCATCACGCTGCTCATCCCCGAGGAGGAGGAGGAGGATGGCTCTATGGAAGAACGACGTG 1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1372 GTCATCACGCTGCTCATCCCCGAGGAGGAGAAGGGATGGCTCGGTCCTATGGAAGAACGACGTG 1431

Qy   1372 TCCAAGGCGGAGGGGTTGGTTCCGTGCGTGGTCACGGAAGTTGCTGGAAGAAAATGAGACA 1431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1372 TCCAAGGCGGAGGGGTTGGTTCCGTGCGTGGTCACGGAAGTTGCTGGAAGAAATGAGACA 1431
```

```
Qy    232 GAGGAGGTGAACCGGGCTCACGGAGAGCACCTACCGGAAGAGCAACTCCCTCGTAATTAGGATCGCCCAGAATGTTATGGAACAGTTCAATCCT 291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 GTGGGGGGTGAAACTCCCGTCAGAATGTTATGGAACAGTTCAATCCT 351
```

```
Db     1432  TCCAAGGCGGAGGGGTTGGTTCCCTTCGTCGTACGGAGAGTTGCTGGAAGAAAATGAGACA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1432  GAAGCAGTGACCGTGCCCACGCCCAAGCCCACCAGTGAGAAGCATCAG  1481
             ||||
Db     1492  GAAGCAGTGACCGTGCCCACGCCCAAGCCCACCAGTGAGAAGCTTCAG  1541

RESULT 6
ADV43099
XX
ID     ADV43099 standard; cDNA; 1221 BP.
XX
AC     ADV43099;
XX
DT     10-MAR-2005  (first entry)
XX
DE     Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 727.
XX
KW     microarray; psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 727;
KW     non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW     cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS     Homo sapiens.
XX
PN     WO200410899-A2.
XX
PD     16-DEC-2004.
XX
PF     04-JUN-2004; 2004WO-US017686.
XX
PR     04-JUN-2003; 2003US-0475915P.
XX
PA     (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI     Nicholson A, Vernon SD;
XX
DR     WPI; 2005-031682/03.
XX
PT     New microarray comprising probes for genes involved in
PT     psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT     condition associated with PNI activity, e.g., inflammatory or infectious
PT     diseases.
XX
PS     Claim 1; SEQ ID NO 727; 254pp; English.
XX
CC     The invention relates to a new microarray which comprises probes for
CC     genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC     microarray is useful in diagnosing a condition associated with PNI
CC     activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC     cancer and infection. The present sequence represents a
CC     psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC     specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
CC     SEQ ID NO 1829 are provided.
XX
SQ     Sequence 1221 BP; 397 A; 279 C; 288 G; 257 T; 0 U; 0 Other;

       Query Match             57.6%;  Score 1206.4;  DB 14;  Length 1221;
       Best Local Similarity   99.9%;  Pred. No. 1.9e-198;
       Matches 1207;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy      274  ATGGAACAGTTCAATCCTGGGCTGCGAATTTAATAAACCTGGGGAAAAATTATGAGAA  333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1  ATGGAACAGTTCAATCCTGGGCTGCGAATTTAATAAACCTGGGGAAAAATTATGAGAAA  60

Qy      334  GCTGTAAACGCTATGATCCTTGGCAGGAAAAGCCTACTACCATGGAGTGGCCAAGATCGGT  393
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61  GCTGTAAACGCTATGATCCTTGGCAGGAAAAGCCTACTACCATGGAGTGGCCAAGATCGGT  120

Qy      394  GAGARTTGCCACTGGGTCCCCGTGTCAACTGAACTGGGACATGTCCTCATTAGAGAATTTCA  453
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121  GAGARTGCCACTGGGTCCCCGTGTCAACTGAACTGGGACATGTCCTCATTAGAGAATTTCA  180

Qy      454  AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTAAAAAATTCCACAAAGAG  513
```

```
Db      181  AGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAATTTAAAAAATTCCACAACTCTA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      514  ATTATCCATGAGCTGGAGAAGCAAGATTAGAAGATTAGAACTTGAGGTGAAATATGAACCA  573
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241  ATTATCCATGAGCTGGAGAAGCAAGATTAGAAGATTAGAACTTGAGGTGAAATATGAACCA  300

Qy      574  AAAAGATACCAAAACAGAACACAAGAAGAATAAATTAGAGTCTTTTGGAGAAATCCAAGCTGAG  633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301  AAAAGATACCAAAACAGAACACAAGAAGAATAAATTAGAGTCTTTTGGAGAAATCCAAGCTGAG  360

Qy      634  TTGAAGAAGATCAGAGAGAAAAGCCAAGGAACCGAAAACGCACTCAAATATGAACACAAA  693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361  TTGAAGAAGATCAGAGAGAAAAGCCAAGGAACCGAAAACGCACTCAAATATGAACACAAA  420

Qy      694  GAAATTCGAGTATGTGGAGACACCGTTACTTCTCCGTCAGAGTGTGAAAATCCAGAAATCATTGCA  753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421  GAAATTCGAGTATGTGGAGACACCGTTACTTCTCCGTCAGAGTGTGAAAATCCAGAAATCATTGCA  480

Qy      754  GATGGTTGCAAAGACGGCTCTGCTTGAAGGAAGAGGGCGCCTTCTTGCTTTTCTTCTGTTTGATGAG  813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481  GATGGTTGCAAAGAGGGCTCTCGCTTGAAGGAAGAGGGCGCCTTCTTGCTTTTCTTCTGTTTGATTCCA  540

Qy      814  CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCGAGAACTACTGAAT  873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541  CACTGTGGCTTTGCAAACCACATACATTATTATCACTTACAGTCTGCGAGAACTACTGAAT  600

Qy      874  TCCAAGCTGCCTCGGTTGCAGGAGGAGACCTGTGTTGATGCCATCAAAAGTGCCAGAAAATC  933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601  TCCAAGCTGCCTCGGTTGCAGGAGGAGACCTGTGTTGATGCCATCAAAAGTGCCAGAAAATC  660

Qy      934  ATGAATATGAGTCGAAGAAATAAAGAACCCCAGCCTTCTACCCCCCCTGTCTGGAACTCCTCAG  993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661  ATGAATATGAGTCGAAGAAATAAAGAACCCCAGCCTTCTACCCCCCCTGTCTGGAACTCCTCAG  720

Qy      994  GCTTCACCCATGATCGAGAGAGGCAAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA  1053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721  GCTTCACCCATGATCGAGAGAGGCAAATGTGGTTAGGAAAGATTACGACACCCTTTCTAAA  780

Qy     1054  TGCTCACCAAAGATGCCCCCGCGTCCTTCAGGCAGCAGCATATACCAGTCCCTTGATCGAT  1113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781  TGCTCACCAAAGATGCCCCCGCGTCCTTCAGGCAGCAGCATATACCAGTCCCTTGATCGAT  840

Qy     1114  ATGTTAATAACCCAGCCACAGCGGCTGCCCGGAATTCACAAAAGGGTAAATAATTCAACAGGT  1173
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841  ATGTTAATAACCCAGCCACAGCGGCTGCCCGGAATTCACAAAAGGGTAAATAATTCAACAGGT  900

Qy     1174  ACTTCCGAAGATCCCAGTTTACACGGATCAGTTTGCGGTTGCAACGGGACTGAACATGATG  1233
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901  ACTTCCGAAGATCCCAGTTTACACGGATCAGTTTGCGGTTGCAACGGGACTGAACATGATG  960

Qy     1234  AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC  1293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961  AAGAAGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCTCCAACAAGACCTTACTC  1020

Qy     1294  AGCTTTGCACAGGAGGATGTCATCACGCTGCTCATCCCCGAGGACAAGGAGATGGCTGGCTC  1353
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021  AGCTTTGCACAGGAGGATGTCATCACGCTGCTCATCCCCGAGGACAAGGAGATGGCTGGCTC  1080

Qy     1354  TATGGAGAACACGACGGTGTCCAAGGCCGAGGGGTTGGTTCCCGGTGCTCGTACGGAAGTTG  1413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081  TATGGAGAACACGACGGTGTCCAAGGCCGAGGGGTTGGTTCCCGGTGCTCGTACGGAAGTTG  1140

Qy     1414  CTGGAAGAAAATGAGACAGCAGGAGACCAGTGACCGTGCCCACGCCCAAGCCCACCAGTGAGA  1473
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141  CTGGAAGAAAATGAGACAGCAGTGACACAGGAGACCAGTGACCGTGCCCACGCCCAACACCAGTGAGA  1200

Qy     1474  AGCATCAG  1481
             ||||||
Db     1201  AGCTTCAG  1208

RESULT 7
```

```
AAS73957
ID   AAS73957 standard; cDNA; 1729 BP.
XX
AC   AAS73957;
XX
DT   13-FEB-2002 (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9761.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09770.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 9761; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (II) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1729 BP; 426 C; 389 G; 397 T; 0 U; 0 Other;
```

```
Query Match             51.1%;  Score 1071;  DB 5;  Length 1729;
Best Local Similarity   95.1%;  Pred. No. 4.1e-175;
Matches  1205;  Conservative    0;  Mismatches   45;  Indels   17;  Gaps   9;

Qy   232 GAGGAGGTGAACCGGCTCACGGAGGAGCACCTACGGAGGAATGTTATCGAACAGTTCAATCCT 291
              ||  |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   277 GTGGGGGTGAAACTCCCCTGTAATTAGGATCGCCAGAANGTTATCGAACAGTTCAATCCT 336

Qy   292 GGGCTGCGAAATTTAATAAACCTGGGGAAAAAATTATGAGAAAAGCTGTAAACGCTATGATC 351
         |||||||||||||||||||||||||||||||||| ||||||||| |||||||||||||||||
Db   337 GGGCTGCGAAATTTAATAAACCTGGGGAAAAAATTTATGAGAAAGCTGTAAACGCTATGATC 396

Qy   352 CTGGCAGGAGAAAGCCTACTACGGATGGAGTGGGCCAAGATCGCGGTGGAGATTGCCACTCGGTCC 411
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   397 CTGGCAGGAGAAAGCCTACTACGGATGGAGTGGGCCAAGATCGCGGTGGAGATTGCCACTCGGTCC 456

Qy   412 CCCGGTGTCAAC-TGAACTGGGACATGTCCTCATAGAGATTCAAGTACCCACAAGAAAC- 469
         |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   457 CCCGTGTCAACTGTAACTGGGACATGTCCTCATAGAGATTCAAGTACCCACAAGAAACT 516

Qy   470 TCAACCAGAGTCTT----DATGAAAAATTTTAAAAAAATTCCACAAAAGAGATTATCCCATGAGC 526
         |||||||||||||     |||||||||||||||||||||||||||||||||||||||||||||||
Db   517 TCAACCAGAGTCTTTGATGGAAAATTTTTAAAAAAATTCCACAAAAGAGATTATCCCATGAGC 576

Qy   527 TGGAGAAGAAGATAGAACTT-GACGTTGAAAATAATATGAACGCAAACTCTAAAAGAATACCAA 585
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   577 TGGAGAAGAAGATAGAACTTCGACGTTGAAAATATATGAACGCAAACTCTAAAAGAATACCAA 636

Qy   586 ACAGAACACAACAAGAATAAAATTAGAGTGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAGATC 645
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   637 ACAGAACACAACAAGAATAAAATTAGAGTGTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAGATC 696

Qy   646 AGAAGGAAAAAGCCAAGGAAGCCGAAAACGCACTCAAAATATGAACACAAAGAAAATTGACTAT 705
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   697 AGAAGGAAAAAGCCAAGGAAGCCGAAAACGCACTCAAAATATGAACACAAAGAAAATTGAGTAT 756

Qy   706 GTGGGACACCGTTACTTCTCGTCGAGACTGAAAATCC-AGAAAATTCAATTGCAGATGGTTGCAA 764
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||| 
Db   757 GTGGGACACCGTTACTTCTCGTCGAGAGTGAAAATCCTAGAAAATCCATTGCAGATGGTTGCAA 816

Qy   765 AGAGGCTCTGCTTGAAGAGAAGAAGGGCG-------CTTCTGCTTTCTGGTTGATAAGCACTG 818
         ||||||||||||||||||||||||||||        |||||||||||||||||||||||||| 
Db   817 AGAGGCTCTGCTTGAAGAGAAGAGGCGCTTCCTGCTTTCTGGGTTGATGAATAAGCACTG 876

Qy   819 TGGGCTTTGCAAA--CCACATACATATTATTATCACTTACAGTCTGCAGAACTACTGAATT-C 875
         |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||| |
Db   877 TGGGCTTTGCAAAACCCACATACATTATTATCACTTACAGTCTGCAGAACTACTGAATTCC 936

Qy   876 CAAGCTGCCTCGGGTCGGCAGGAGAGAGCCTGTTGTTGATAGTGGCCATCAAAAGTG-CCAGAGAAAATCA 934
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| 
Db   937 CAAGCTGCCTCGGGTCGGCAGGAGAGCCTGTTGTTGATGGCCATCAAAAGTGCCCAGAGAAAATCA 996

Qy   935 TGAATATGATCGAAGAGAAATAAAGACCCCAGCCTCTRCCCCGTGTCTCGGAACTCCTCAGG 994
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   997 TGAATATGATCGAAGAGAAATAAAGACCCCAGCCTCTACCCCGTGTCTCGGAACTCCTCAGG 1056

Qy   995 CTTCACCCATGATCGAGAGAGGAAGCAATGTGGTTAGGAAAAGATTACGGACACCCTTTCTAAAT 1054
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1057 CTTCACCCATGATCGAGAGAGGAAGCAATGTGGTTAGGAAAAGATTACGGACACCCTTTCTAAAT 1116

Qy  1055 GCTCACCAAAGATNGCCCCCGCTCTTCCTTCAGGCAGAAGCCATATACCAGTCCCCTTGATCGATA 1114
         |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1117 GCTCACCAAAGATGGCCCCCGCTCTTCCTTCAGGCAGAAGCATATACCAGTCCCCTTGATCGATA 1176

Qy  1115 TGTTTAATAACCCAGCCACGGGCTGCCCCGGAATTCACAAAAGGGTAAATAAATTAATTCAACAGGTA 1174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1177 TGTTTAATAACCCAGCCACGGGCTGCCCCGGAATTCACAAAGGGTAAATAAATTCAACAGGTA 1236

Qy  1175 CTTCCGAAGATCCCAGTTTACAGCGGATCAGTTTCCGGTTGCAAACGGGACTGAACATGATGA 1234
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1237 CTTCCGAAGATCCCAGTTTACAGCGGATCAGTTTCCGGTTGCAAACGGGACTGAACATGATGA 1296

Qy  1235 AGAAGCAGAAAGTGAAGACCATCTCTCCCGCACACTGCGGGCTCCAACACAGACCTTACTCA 1294
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1297 AGAAGCAGAAAGTGAAGACCATCTCTCCCGCACACTGCGGGCTCCAACACAGACCTTACTCA 1356

Qy  1295 GCTTTGCACAGGGAGAGTGTCATCACGGCTGCTCATCACCCCCGAGGAGGAAGGAATGGCTCGGCTCT 1354
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1357 GCTTTGCACAGGGAGAGTGTCATCAGCTGCTCATCCCCGAGGAGGAAGGAATGGCTCGGCTCT 1416

Qy  1355 ATGGAGAACACGACGTTCCAAGGCGGAGGGTTGGTTCCGGTCGTCGTACGAAGTTGC 1414
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1417 ATGGAGAACACGACGTTCCAAGGCGGAGGGTTGGTTCCGGTCGTCGTACGAAGTTGC 1476

Qy  1415 TGGGAAGAAATGAGACAGAAGCAGTGACCGTGCCCAAGCCCACACCAGTGAGAA 1474
```

```
Db   1477  TGGAAGAAAAATGAGACAAGAAGCAGGTGACCGGTGCCCACGCGCAGCCCAAGCCCACACCAGTGAGAA  1536
                                                                                
Qy   1475  GCATCAG  1481
           |||||||
Db   1537  GCTTCAG  1543
```

RESULT 8
AAS73958

ID   AAS73958 standard; cDNA; 1993 BP.
XX
AC   AAS73958;
XX
DT   13-FEB-2002 (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9762.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09771.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 9762; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (III) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   genes. (I) and products of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1993 BP; 542 A; 491 C; 472 G; 487 T; 0 U; 1 Other;

Query Match   48.7%; Score 1021.2; DB 5; Length 1993;
Best Local Similarity   89.3%; Pred. No. 1.6e-166;
Matches 1211; Conservative 0; Mismatches 38; Indels 107; Gaps 6;
```
Qy    680  AATATGAACAACAAAGAAAATTGAGTATGTGGGAGACGGGTTACTTCTTCGTCAGAGTGAAATCC  739
Db    158  AACTTGACGTGAAATATTATGAAACTATTGTTGGAGAACGTAGAGGAGAAGAGGCGGTTACTTCGTCAGAGTGAAATCC  217

Qy    740  AGAAATTCATTGCAGATGGTTGCAAAGAGGCTTCGCTTGAAGAAGAGGAGGACCGTTACTTCTTCGTCTG  799
Db    278  TTCTGGTTGATTAAGACACTGTTGCAAAACCACATTATTATCACTTACAGTCTG  337

Qy    800  CAGAACTACTGAAATTCCAAGCTGCCTTCGGTGGCAGGAGACCTGTTGTTGATGGCCATCAAAG  859
Db    338  CAGAACTACTGAAATTCCAAGCTGCCTCGGTGGCAGGAGACCTGTTGTTGATGGCCATCAAAG  397

Qy    860  TGCCAGAGAAAATCATGAATAATGAGATGATCGAAGAAAATAAAGACCCCAGCCTTACCCCGTGT  919
Db    398  TGCCAGAGAAAATCATGAATAATGAGATGATCGAAGAAAATAAAGACCCCAGCCTTACCCCGTGT  457

Qy    920  CTGGAACTCCCTCAGGCTTCACCCATGATCGAGAGAGGAAGCAATGTGGTTAGGAAAAGATTACG  979
Db    458  CTGGAACTCCTTCACCCATGATCAGACGAGAAGCAATGTG-----  501

Qy    980  CTGGAACTCGTTCAGCCTTCACCCATGATCGAGAGAGGAAGCAATTACG  1039
Db    502  ----TCCCTGTTCCTTCTCTG---  517

Qy   1040  ACACCCTTTCTAAATGCTCACCAAAGATGCCCCCGCGCCTCCTTCAGGCAGGCATATACCA  1099
Db    518  ------GCAATGATCAAAAGCGAGG  535

Qy   1100  GTCCCTTGATCGATATGTTTAATAACCCAGCCACGGCGTGCCCCGAATTCACAAAGGGTAA  1159
Db    536  GACTGAACATGATGAGAAGCAGAAGTGAAGACAGCATCTTCCCGGACACTGCGGGGCTCCA  595

Qy   1160  ATAATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGGATCAGTTTCGGTTGCAACGG  1219
Db    596  GACTGAACATGATGAGAAGCAGAAGTGAAGACAGCATCTTCCCGGACACTGCGGGGCTCCA  655

Qy   1220  GACTGAACATGATGAGAAGCAGAAGTGAAGACCAATCTTCCCGGACACTGCGGGGCTCCA  1279
Db    656  ACAAGACCTTACTTCAGCTTTGCACAGGGAGGATGTCATCCCGGAGGA  715

Qy   1280  ACAAGACCTTACTTCAGCTTTGCACAGGGAGGATGTCATCCCGGAGGA  1339
Db    716  AGGATGGCTGGCTTCTATGGAGAACACGACGTGTTCCCCGTCGT  775

Qy   1340  AGGATGGCTGGCTTCTATGGAGAACACGACGTGTTCCCCGTCGT  1399
Db    776  CGTACACGGAGTTGCTGGGAGGAAAGAGAGACAGAGTCGACCGTGCCACGCCAAGCC  835

Qy   1400  CGTACACGGAGTTGCTGGGAGGAAAGAGAGACAGAGTCGACCGTGCCACGCCAAGCC  1459
Db    836  CCACACCAGTGAGGAGCATCAGCCGTGCAACTTGTCTGAGAATAGCAGTGTTGTCATCC  895

Qy   1460  CCACACCAGTGAGGAAGCATCAGCCGTGGAGACTTGTCTGAGAATAGCAGTGTTGTCATCC  1519
Db    896  CCCCCACCCGACTACTTGAAATGCTTGTCCATGGGGAGGCTGCCGGACAGGAGGACAGATT  955

Qy   1520  CCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGAGGCTGCCGGACAGGAGGACAGATT  1579
Db    956  CGGCCCAGGACGACATCCACCTTTAAGGCCCCAGCCGTCCAAAGCGGAGGGAGGAGCTCCTA  1015

Qy   1580  CGGCCCAGGACGGCATCCACCTTTAAGGCCCCCAGCCGTCCAAGGCGCGGGGGCTCCTA  1639
Db   1016  ACGATGCCAACGGACTGCAAACCGGCCTTTTCTCAGGGGAGAAAAACCCCTTTGCCACTG  1075

Qy   1640  ACGATGCCAACGGGGACTTGCAAACCGGCCTTTTTCTCAGGGGAGAAAAACCCCTTTGCCACTG  1699
Db   1076  TGAAACTCCGCCCGGAATGATCGGCTGGCAGCCCATCATTCGGATTGACGAGAGGACA  1135

Qy   1700  TGAAACTCCGCCCGGAAATGATCGGCTGGCAGCCCATCATTCGGATGACGAGAGGACA  1759

```
Qy  1760  GCCAAGGAGACTCTCCC--GGGGCCTCTCCCGGGTTCTCCCTTGCGGAATGATGGGGC---GCATCCT  1816
          ||||||||||||||||   |||||||||||| |||||||||||||||||||||||||    |||||| 
Db  1136  GCCAAGGAGACTCTCCCAGGGGCCTCTCCCGGTTCTCCCTTGCGGAATGATGGGCCGCATCCCT  1195

Qy  1817  GTCTGCCACGTGCTGACGGTC-GGGAAGCTRCAGTGGAGAGGGAGAGGGAGGGCCTRACRCTCTAAGTCGGCC  1875
          ||||||||||||||||||||| |||||||| |||||| |||||||   ||||||||  ||| ||||||||||| 
Db  1196  GTCTGCCACGTGCTGACGGTCGGGGAAGCTTCAGTGGGAGAGGGAGGGCCTRAACTCTAAGTCGGCC  1255

Qy  1876  TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTYGAC-AAGTTTCTGCCTTT  1934
          ||||||||||||||||||||||||||||||||||||||||| |||  |||||||||||||
Db  1256  TGCTTAAGCAAATCATGCTTCTCTGTTTCACGTAGTTGGGTTGACAAAGTTTCTGCCTTT  1315

Qy  1935  AAGATAAAATGAGTAATAGTTGTCTTAATCACCAGCTCACGCCCATTTAAATATATTTTTCTTCCTTT  1994
          |||||||||||||||||||||  |||||||||||||||| |||||||||| |||||||||||||| ||
Db  1316  AAGATAAAATGAGTAATAGTTCGAATCACCAGCTCAGCCATTTAAAATATATTTTTCTTCCTATT  1375

Qy  1995  CTGTTCAAGAAACACAGTAAAACTTGGTTGTTTCAATCTTTA  2030
          ||||||||||||||| |||||||||||||| |||||||||||
Db  1376  CTGTTCAAGAAACAGTAAAACTTGGTTTCAATCTTTA  1411
```

RESULT 9

```
ID   ADT94962 standard; cDNA; 625 BP.
XX
AC   ADT94962;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #481.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-0066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313007P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
PS   Claim 1; SEQ ID NO 481; 87pp; English.
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
SQ   Sequence 625 BP; 215 A; 127 C; 159 G; 124 T; 0 U; 0 Other;
```

Query Match   28.3%;   Score 594;   DB 11;   Length 625;
Best Local Similarity   100.0%;   Pred. No. 5.3e-93;
Matches 594;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy  205  GCGGCCGCAGCCATGTCCCCGGCCCGAGGAGGGTGAACCGGCTCACGCGAGGAGCACCTAC  264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32  GCGGCCGCAGCCATGTCCCCGGCCCGAGGAGGGTGAACCGGCTCACGCGAGGAGCACCTAC   91

Qy  265  CGGAATGTTATGGAACAGTTCAATCCTGGCTGCCGGAATTTAATAAACCTGGGGGAAAAT  324
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   92  CGGAATGTTATGGAACAGTTCAATCCTGGCTGCCGGAATTTAATAAACCTGGGGGAAAAT  151

Qy  325  TATGAGAAAGCTGTRAAACGCTATAGATCCTGGCCAGGAAAAGCCTACTACGATGGAGTGGGCC  384
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  152  TATGAGAAAGCTGTRAAACGCTATAGATCCTGGCCAGGAAAAGCCTACTACGATGGAGTGGGCC  211

Qy  385  AAGATCGGTGAGATTGCCACTGGGGTCCCCCCGTGTCAAACTGGGACATGTCCTCATA  444
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  212  AAGATCGGTGAGATTGCCACTGGGGTCCCCCCGTGTCAAACTGGGACATGTCCTCATA  271

Qy  445  GAGATTTCAAGTACCCACAAGAAAACTCAACGAGAGACTCTTGATGAAAAATTTTAAAAAATTC  504
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  272  GAGATTTCAAGTACCCACAAGAAAACTCAACGAGAGACTCTTGATGAAAAATTTTAAAAAATTC  331

Qy  505  CACAAAGAGAGNATTATCCATGAGCTGGAGAAGAAGAGATTAGAGACTTGACGTGAAATATATGAAC  564
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  332  CACAAAGAGAGNATTATCCATGAGCTGGAGAAGAAGAGATTAGAGACTTGACGTGAAATATATGAAC  391

Qy  565  GCAACTCTAAAAAGATACCAAACACAGAACAACAAGAGAATAAATTAGAGTCTTTGGAGAAATCC  624
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  392  GCAACTCTAAAAAGATACCAAACACAGAACAACAAGAGAATAAATTAGAGTCTTTGGAGAAATCC  451

Qy  625  CAAGCTGAGTTTGAAGAAGATCAGAAGGAAAAAGCCAAGGAAACGACTCAAATAT  684
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  452  CAAGCTGAGTTTGAAGAAGATCAGAAGGAAAAAGCCAAGGAAACGACTCAAATAT  511

Qy  685  GAACACAAAGAAAATTGAGTRATGTGGAGGACCGGTTACTTCTCCGTCAGAGTGAAATCCAGAAA  744
         ||||||||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||
Db  512  GAACACAAAGAAAATTGAGTRATGTGGAGGACCGGTTACTTCTCGTCAGAGTGAAATCCAGAAA  571

Qy  745  TTCATTGCGATGGTTGCAAAGAGGCTTCTGCTTGAAGAGAGAAGCGCGCTTCTGC  798
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  572  TTCATTGCGATGGTTGCAAAGAGGCTTCTGCTTGAAGAGAGAAGCGCGCTTCTGC  625
```

RESULT 10

```
ID   ADX41444 standard; cDNA; 625 BP.
XX
AC   ADX41444;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 481.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   88; gene.
```

```
XX
OS   Homo sapiens.
XX
PN   WO20274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 481; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 215 A; 127 C; 159 G; 124 T; 0 U; 0 Other;

Query Match             594; Score 594; DB 11; Length 625;
Best Local Similarity   100.0%;
Matches  594; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   205 GCGGCCCCAGCCACATGTCCCGGGGGGCCCCGAGGAGTGAACCGGCCACCTAC 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    32 GCGGCCCGCAGCCATGTCCCGGGGGGCCCCGAGGAGTGAACCGGCCACCTAC 91

Qy   265 CGGAAATGTTATGGAACACGTTCAATCCTGGGCTCGGAAATTTAATAAAACCTGGGGAAAAT 324
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    92 CGGAAATGTTATGGAACAGTTCAATCCTGGGCTCGGAAATTTAATAAAACCTGGGGAAAAT 151

Qy   325 TATGAGAAGAAGCTGTAAAACGCTATGATCCTGGCGAGGAAAAGCCTACTACGGATGGAGTGGCC 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 TATGAGAAGAAGCTGTAAAACCCTATGATCCTGGCGAGGAAAAGCCTACTACGGATGGAGTGGCC 211

Qy   385 AAGATCGGTGAGATTGCCACTGGGGTCCCCCGTGTCAACTGAACTGGGACCATGTCCTCATA 444
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212 AAGATCGGTGAGATTGCCACTGGGTCCCCCGTGTCAACTGAACTGGGACCATGTCCTCATA 271

Qy   445 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAAATTTTAAAAAATTC 504
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   272 GAGATTTCAAGTACCCACAAGAAACTCAACGAGAGTCTTGATGAAAAATTTTAAAAAATTC 331

Qy   505 CACAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGGTAGAAATATATGGAAC 564
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   332 CACAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGGTAGAAATATATGAAC 391

Qy   565 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGGAGAAAATCC 624
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   392 GCAACTCTAAAAAGATACCAAACAGAACACAAGAATAAATTAGAGTCTTTGGGAGAAATCC 451

Qy   625 CAAGCTGAGTTGAAGAAGGATCAGAAGGAAAAAGCCAAGGAGAAGCCGAAGCGCACTCAAATAT 684
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 CAAGCTGAGTTGAAGAAGGATCAGAAGGAAAAAGCCAAGGAGAAGCCGAAGCGCACTCAAATAT 511

Qy   685 GAACACAAAAGAAATTGAGTATGTGGAGACCGGTTACTTCGTCAGAGTGAAATCCAGAA 744
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   512 GAACACAAAAGAAATTGAGTATGTGGAGACCGGTTACTTCGTCAGAGTGAAATCCAGAA 571

Qy   745 TTCATTGCGAGATGGTTGCAAAGAGGCCTTCGCTTGACGAGAGAGAGAGGGCCGCTTCTGC 798
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   572 TTCATTGCGAGATGGTTGCAAAGAGGCCTTCGCTTGACGAGAGAGAGAGGGCCGCTTCTGC 625

RESULT 11
AC   AAC77875
ID   AAC77875 standard; cDNA; 1242 BP.
XX
AC   AAC77875;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:269.
XX
PI   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO20055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005882.
XX
PR   12-MAR-1999;  99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43666.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer.
XX
PS   Claim 1; Page 832-833; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC   AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemolytic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
```

```
CC    bacterial or viral infections. The peptides, nucleotides, antibodies,
CC    agonists and antagonists may be also be used in drug screens. AAC78449 to
CC    AAC78457 and AAB44240 represent sequences used in the exemplification of
CC    the present invention
XX
SQ    Sequence 1242 BP; 322 A; 291 C; 281 G; 338 T; 0 U; 10 Other;

      Query Match             27.5%;  Score 576.4;  DB 3;  Length 1242;
      Best Local Similarity   95.3%;  Pred. No. 6.2e-90;
      Matches 592;  Conservative 1;  Mismatches 28;  Indels 0;  Gaps 0;

Qy  1410 GTTGCTGGAAGAGAAAATGAGACAGAAGCAGTGACCGTGCCCAGCCCACACACAGT 1469
         |||||||||||||||||
Db   239 GGTGGTGAGAGACATTATCTAACACGTGCGTGTCCTTTTGGGTTTGAGCCCGACACACAGT 298

Qy  1470 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCGCCCCCGA 1529
         ||||||||||||||||||||||||||||||||||||||||||||
Db   299 GAGAAGCATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCACCCCCGA 358

Qy  1530 CTACTTGGAAATGCTTGTCCATGGGGGCCGCAGCAGGAGGAGCAGAATTCGGCGTAGGAC 1589
         |||||||||||||||||||||||||||||||||||||||||||
Db   359 CTACTTGGAAATGCTTATTCCATGGGGGGCCAGYGTCCCAGCAGGAGACAGATTCGGCCAGGAC 418

Qy  1590 GACATCCACCTTTAAGGCCCCAGGGTCCAAGCCCGGGGCTCCTAACGATCCAA 1649
         |||||||||||||||||||||||||||||||||||||||||
Db   419 GACATCCACCTTTAAGGCCCCAGGCGTCCAAGCCCGGAGACCGNGGGCTCCTAACGATGCCAA 478

Qy  1650 CGGGACTGCAAAGCCCGCCTTTTCTCAGCGGCGGAGAAAACCCCTTTGCCACTGTGAACTCCG 1709
         ||||||||||||||||||||||||||||||||||||||||
Db   479 CGGGACTGCAAAGCCCGCCTTTTCTCAGCGGCGAGAAAACCCCTTTGCCACTGTGAAACTCCG 538

Qy  1710 CCCGACTGTGACGAATGATGAATGCGCTCGGCACCCATCATTCGATGAGAGGAGGACAGCCAAGGACT 1769
         |||||||||||||||||||||||||||||||||||||||
Db   539 CCCGACTGTGACGAATGATCGCTCGGCACCCATCATTCGATGAGAGGAGGACAGCCAAGGACT 598

Qy  1770 CTCCCGGGGCCTCTCCGGTTCTCCCTTGCGGAATGATGGGGCGGGCCATCGTCTCGTCTGTGCCACGTGC 1829
         |||||||||||||||||||||||||||||||||||
Db   599 CTCCCGGGGCCCTCTCCGGTTCTCCCTTGCGGAATGGGGCGCCATCCGTCTCCGTGCTCGCTCCACGTGC 658

Qy  1830 TGACGGTCCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAACTGCTTAAGCAAAATC 1889
         ||||||||||||||||||||||||||||||||||||||||||
Db   659 TGACGGTCCGGGAAGCTTCAGTGGAGAGGCCTAACTCTAAATGTGCGCCTGCTTAAGCAAAATC 718

Qy  1890 ATGCTTCTCTGTTTCACGTAGTTCGGGTTCGGCTTGACAAGTTCTCTCGTTTTAAGATAAATGAGTAA 1949
         |||||||||||||||||||||||||||||||||||||||
Db   719 ATGCTTCTCTGTTTCACGTAGTTCGGGTTGACAAGTTCTCGCCTTTAAGAATAAATGAGTAA 778

Qy  1950 TAGTCTAATGACCAGCTCAGCCATTTAAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 2009
         ||||||||||||||||||||||||||||||||||||
Db   779 TAGTCTAATGACCAGCTCAGCCATTTAAAAATATTTTCTTCCTATTCTGTTCAAGAAACAG 838

Qy  2010 TAAACTTGGTTTCAATCTTTA 2030
         ||||||||||||||||||||
Db   839 TAAACTTGGTTTCAATCTTTA 859
```

```
RESULT 12

AC    ADT95471;
ID    ADT95471 standard; cDNA; 641 BP.
XX
DT    16-DEC-2004  (first entry)
DE    Colon cancer associated human cDNA sequence #990.
KW    Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW    humoral immune response; cellular immune response; cytostatic;
KW    immunostimulant; human; ss.
OS    Homo sapiens.
XX
PN    US20030878818-A1.
```

```
XX
PD    08-MAY-2003.
XX
PF    01-FEB-2002; 2002US-00065543.
XX
PR    02-FEB-2001; 2001US-0267400P.
PR    07-FEB-2001; 2001US-0267382P.
PR    11-MAY-2001; 2001US-0290322P.
PR    12-JUL-2001; 2001US-0305265P.
PR    16-AUG-2001; 2001US-0313077P.
XX
PA    (CORI-) CORIXA CORP.
XX
PI    Jiang Y, Chenault RA, Xu J, Indilas CY, Lodes MJ, Secrist H;
PI    Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR    WPI; 2003-040540/03.
XX
PT    New isolated nucleic acids and polypeptides capable of eliciting a
PT    humoral and/or cellular immune response, useful for diagnosing,
PT    preventing or treating cancer, particularly colon cancer.
XX
PS    Claim 1; SEQ ID NO 990; 87pp; English.
XX
CC    The invention relates to polynucleotide and polypeptide sequences
CC    associated with cancer, particularly colon cancer. Also disclosed are (i)
CC    an expression vector comprising the polynucleotide, (ii) a host cell
CC    transformed or transfected with the expression vector, (iii) an isolated
CC    antibody, or its antigen-binding fragment, which specifically binds to
CC    the polypeptide, (iv) a method of detecting or determining the presence
CC    of cancer in a patient, (v) a fusion protein comprising at least one of
CC    the polypeptides, (vi) an oligonucleotide that hybridises to the
CC    polynucleotide sequence under highly stringent conditions, and (vii) a
CC    method of stimulating and/or expanding T cells specific for a tumour
CC    protein. The polypeptide specifically comprises the amino acid sequence
CC    of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC    frames (ORFs) of C636S. These polypeptides are encoded by the
CC    polynucleotide sequences, where both are capable of eliciting a humoral
CC    and/or cellular immune response. The polynucleotides, polypeptides, and
CC    antibodies are useful for diagnosing, preventing or treating cancer,
CC    particularly colon cancer. The polynucleotide and polypeptide sequences
CC    are also useful in DNA strand invasion, antisense inhibition, mutational
CC    analysis, nucleic acid purification, isolation of transcriptionally
CC    active genes, blocking or transcription factor binding, genome cleavage
CC    or in situ hybridisation, and as enhancers of transcription or
CC    biomarkers. This sequence represents a human colon cancer associated
CC    cDNA. Note: The sequence data for this patent was obtained in electronic
CC    format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ    Sequence 641 BP; 217 A; 127 C; 154 G; 129 T; 0 U; 14 Other;

      Query Match             27.5%;  Score 576.2;  DB 11;  Length 641;
      Best Local Similarity   97.3%;  Pred. No. 6.2e-90;
      Matches 578;  Conservative 0;  Mismatches 16;  Indels 0;  Gaps 0;

Qy   205 GCGGCGCGCAGCCATGTCCCGGCGGGCCCGAGGAGGTGAACCGGCTCACGGAGGAGCACTAC 264
         |||||||||||||||||||||||||||||||
Db    48 GCGGCGCCGCAGCCATGTCCCGGCGGGCCCGAGGAGGTGAACCGGCTCACGGAGGAGCACCTAC 107

Qy   265 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCCAGAATTTAATAAAACCTGGGGAAAAT 324
         |||||||||||||||||||||||||||||||||||
Db   108 CGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGAAAAT 167

Qy   325 TATGAGAAAGCTGTAAACCGCTATATGCTTCCTGGCAGGAGAAAAGCCTACTACGATGGAGCTGGCC 384
         ||||||||||||||||||||||||||||||||||
Db   168 TATGAGAAAGCTGTAAACCGCTATGATCCTTCCTGGCAGGAGAAAAGCCTACTACGATGGAGCTGGCC 227

Qy   385 AAGATCGGTGAGATTGCCACTGGGTCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 444
         |||||||||||||||||||||||||||||||||
Db   228 AAGATCGGTGAGATTGCCACTGGGTCCCCGTGTCAACTGAACTGGGACATGTCCTCATA 287

Qy   445 GAGATTTCAAGTACCCACACAGAAAACTCAACGAGAGTCTTGATGAAAATTTAAAAAATTC 504
         ||||||||||||||||||||||||||||||||||||||||
```

Db   288 GAGATTTCAAGTAGTACCCACAAGAGAAACTCAACGAGAGAGTCTTGATGAAAATTTAAAAATTC 347
          |||||||||||| |||||||||||||||||||||| | |||||||||||||| |||||||
Qy   505 CACAAAGAGAATTATCCCATGAGCTGGAGAAGAGAAGATAGAAGCTTGAGTGAAATATATGAAC 564
          |||||||||||||| |||||||||||||||||||||||||||| ||||| |||||||||||||
Db   348 CACAAAGAGAATTATNCATGAGCTGGAGAAGAGAAGATAGAACTTGACGTGAAATATATGAAC 407

Qy   565 GCAACTCTTAAAAAGATACCAAAACAGAATAAAATTAGAGTCTTTGGAGAAATCC 624
          |||||||||||| |||| || |||||||| | | ||||| | ||||||||||||
Db   408 GCAACTCTTAAAAGATNCCNAACAGAACACACAAGAATNAATTNGAGTCTTTTGAGAAATCC 467

Qy   625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAGAAAACCGAAAACGACTCAAATAT 684
          ||||||||||||||||||||||||||||| |||| |||| |||||||||||||||
Db   468 CAAGCTGAGTTGAAGAAGATCAGAAGGAAGCCCAGAAGCGAAAAACGCACTCAAATAT 527

Qy   685 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTCTTCGTCAGAGTGAAAATCCAGAAA 744
          ||||||||||||||||||||||||||||||||||||||| | |||||||||||| ||||| |||
Db   528 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTNTNGTCAGAGTGAAATCCANAAA 587

Qy   745 TTCATTGCAGATGGGTTGCAAAGAGGCTTCTGCTTGAAGAGAAGAGAGGGCTTCTGC 798
          ||||||||||| ||| |||||| | ||| ||||||||||| || ||| |||||||||
Db   588 TTCATTGCAGAATGGGNTGCAAAANANGCTCTGCTTGAANAGAANANAGGCTTCTGC 641

RESULT 13
ID   ADX41953
AC   ADX41953;
XX
DT   21-APR-2005 (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 990.
XX
KW   88; gene.
XX
OS   Homo sapiens.
XX
PN   WO20274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
PS   Claim 1; SEQ ID NO 990; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.

XX
SQ   Sequence 641 BP; 217 A; 127 C; 154 G; 129 T; 0 U; 14 Other;
     Query Match   27.5%; Score 576.2; DB 11; Length 641;
     Best Local Similarity 97.3%; Pred. No. 6.2e-90;
     Matches 578; Conservative   0; Mismatches 16; Indels   0; Gaps   0;

Qy   205 GCGGCCGCAGCCATGTCCCGGCGGGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGCACCTAC 264
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   48 GCGGCCGCAGCCATGTCCCGGCGGGGGGGGCCCGAGGAGGTGAACCGGCTCACGGAGCACCTAC 107

Qy   265 CGGAATGTTATGGACAGTTCAATCTCTGGGGTTGCCGAAATTTAATAAACCTGGGGAAAAT 324
          |||||||||||||| ||||||||| |||||| |||||||||||||||||||||||||||||
Db   108 CGGAATGTTATGGAACAGTTCAATCCTGGGGCTGCCGAAATTTAATAAACCTGGGGAAAAT 167

Qy   325 TATGAGAAAGCTGTAAACGCTATGAGATCCTGCCAGGAAAAAGCCTACTACGATGGAGTGGCC 384
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   168 TATGAGAAAGCTGTAAAACGCTATGAGATCCTGCCAGGAAAAAGCCTACTACGATGGAGTGGCC 227

Qy   385 AAGATCGGTGACGATTGCCACTGGGGTCCCCCGGTTGTCAACTGAACGTGGGACCATGTCCTCATA 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   228 AAGATCGGTGAGAATTGCCACTGGGGTCCCCCGGTTGTCAACTGAACGTGGGACCATGTCCTCATA 287

Qy   445 GAGATTTCAAGTACCCCACAAGAGAAACTCAACCGAGAAACTCTTGATGAAAATTTTAAAAAATTC 504
          |||||||||||| ||||||||||||||||||| |||| | ||||||||||||| |||||||
Db   288 GAGATTTCAAGTAGTACCCACAAGAGAAACTCAACGAGAGAGTCTTGATGAAAATTTTAAAAATTC 347

Qy   505 CACAAAGAGAATTATCCCATGAGCTGGAGAAGAGAAGATAGAAGCTTGAGTGAAATATATGAAC 564
          |||||||||||||| |||||||||||||||||||||||||||| ||||| |||||||||||||
Db   348 CACAAAGAGAATTATNCATGAGCTGGAGAAGAGAAGATAGAACTTGACGTGAAATATATGAAC 407

Qy   565 GCAACTCTTAAAAAGATACCAAAACAGAATAAAATTAGAGTCTTTGGAGAAATCC 624
          |||||||||||| |||| || |||||||| | | ||||| | ||||||||||||
Db   408 GCAACTCTTAAAAGATNCCNAACAGAACACACAAGAATNAATTNGAGTCTTTTGAGAAATCC 467

Qy   625 CAAGCTGAGTTGAAGAAGATCAGAAGGAAGAAAACCGAAAACGACTCAAATAT 684
          ||||||||||||||||||||||||||||| |||| |||| |||||||||||||||
Db   468 CAAGCTGAGTTGAAGAAGATCAGAAGGAAGCCCAGAAGCGAAAAACGCACTCAAATAT 527

Qy   685 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTCTTCGTCAGAGTGAAAATCCAGAAA 744
          ||||||||||||||||||||||||||||||||||||||| | |||||||||||| ||||| |||
Db   528 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTNTNGTCAGAGTGAAATCCANAAA 587

Qy   745 TTCATTGCAGATGGGTTGCAAAGAGGCTTCTGCTTGAAGAGAAGAGAGGGCTTCTGC 798
          ||||||||||| ||| |||||| | ||| ||||||||||| || ||| |||||||||
Db   588 TTCATTGCAGAATGGGNTGCAAAANANGCTCTGCTTGAANAGAANANAGGCTTCTGC 641

RESULT 14
ID   ABK29726
AC   ABK29726;
XX
DT   23-APR-2002 (first entry)
XX
DE   Colon adenocarcinoma-specific cDNA #252.
XX
KW   Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20196389-A2.
XX
PD   20-DEC-2001.
XX
PF   07-JUN-2001; 2001WO-US018574.
XX
PR   09-JUN-2000; 2000US-0210667P.
PR   22-NOV-2000; 2000US-0256614P.
XX
PA   (CORI-) CORIXA CORP.
XX
XX
ID   ABK29726 standard; cDNA; 559 BP.

PI   Meagher MJ, King GE, Xu J, Secrist H;
XX
WPI; 2002-098052/13.
DR
XX
PT   New isolated polynucleotide encoding a polypeptide comprising a portion
PT   of colon tumour protein, for detection, diagnosis and therapy of human
PT   colon cancer.
XX
PS   Claim 1; Page 178-179; 211pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) encoding a
CC   polypeptide (II) comprising a portion of a colon tumour protein. A new
CC   oligonucleotide (III) that hybridises to (I) is useful for determining
CC   the presence of a cancer in a patient. (II) or antigen presenting cells
CC   expressing (I) are useful for stimulating and/or expanding T cells
CC   specific for a tumour protein, by contacting T cells with (I), (II), or
CC   antigen-presenting cells that express (I), (II), or antigen
CC   presenting cells that express (I), (II), or antigen
CC   in a patient by incubating CD4+ and/or CD8+ T cells isolated from a
CC   patient with (I), (II), or antigen presenting cells that express (II), so
CC   that T cells proliferate, and administering to the patient an effective
CC   amount of the proliferated T cells, thus inhibiting the development of a
CC   cancer in the patient. A new composition is useful for stimulating an
CC   immune response in a patient. (I) or (II) is useful in vaccines and
CC   pharmaceutical compositions for prevention and treatment of colon cancer
CC   and for the diagnosis and monitoring of the cancers. (I), (II) or an
CC   antibody against (II) is useful for detection, diagnosis and/ or therapy
CC   of human colon cancer. (I) is useful as a probe or primer for nucleic
CC   acid hybridisation, and in the design and preparation of ribozyme
CC   molecules for inhibiting expression of (II) in tumour cells. ABK29475-
CC   ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of
CC   the invention.
XX
SQ   Sequence 559 BP; 160 A; 148 C; 132 G; 116 T; 0 U; 3 Other;

Query Match   26.5%;   Score 556;   DB 6;   Length 559;
Best Local Similarity   99.5%;   Pred. No. 1.8e-86;
Matches 556;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy         863 AACTACTGAATTCCAAGCTCGCCTCGGTGGCCAGGAGACCTGTGTTTGATGCCATCAAAGTGC 922
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AACTACTGAATTCCAAGCTCGCCTCGGNGGCCAGGAGACCTGTGTTTGATGCCATCAAAGTGC 60

Qy         923 CAGAGAAAATCATGAAATATGATCGAAGAAATAAAGACCCCAGCCTTCTACCCCGTGTCTCG 982
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CAGAGAAAATCATGAATATGATCGAAGAAATAAAGACCCCAGCCTTCTACCCCGTGTCTG 120

Qy         983 GAACTCCTCAGGCTTCACCCATGATCGAGAGAAGCAAATGTGGTTAGGAAAGATTACGACA 1042
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GAACTCCTCANNCTTCACCCATGATCGAGAGAAGCAAATGTGGTTAGGAAAGATTACGACA 180

Qy        1043 CCCTTTCTAAATGCTCACCAAAGATGGCCCCCGCGTCCTTCAGGCAGCAGCTATATCCAGTC 1102
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCCTTTCTAAATGCTCACCAAAGATGGCCCCCGCGTCCTTCAGGCAGCAGCTATATCCAGTC 240

Qy        1103 CCTTGATCGATATGTTTAATTAAATAACCCAGCCACACGGCCACACGGCCCCGGAATTCACAAAGGGTAAATA 1162
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCTTGATCGATATGTTTAATTAACCCAGCCACACGGCCCCGGAATTCACAAAGGGTAAATA 300

Qy        1163 ATTCAACAGGTACTTCCGAAGATCCCAGTTTACAGCGATCCAGTTTCGGTTGCAACGGGAC 1222
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATTCAACAGGTACTTCCGAAGATCCCAGTTTCACGCGATCCAGTTTCGGTTGCAACGGGAC 360

Qy        1223 TGAACAATGATGAAGAAGCAGAAGTGAAGAAACCATCTTCCCGCACACTGCGGCCTCCAACA 1282
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGAACAATGATGAAGAAGCAGAAAGTGAAGAACCATCTTCCCGCACTGCGGGCTCCAACA 420

Qy        1283 AGACCTTACTCAGCTTTGCACAGGAGATGTCATCACGTGCTCATCCCCGGAGGAGAAGG 1342
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AGACCTTACTCAGCTTTGCACAGGAGATGTCATCACGTGCTCATCCCCGGAGGAGAAGG 480

Qy        1343 ATGGCTGGGCTCTATGGAGAACAACGACGTGTCCAAGGCGAGGGGGTTGGGTTCCGTCGTCGT 1402
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 15
Db
XX
ID   ADT94802   standard; cDNA; 539 BP.
XX
AC   ADT94802;
XX
DT   16-DEC-2004   (first entry)
XX
DE   Colon cancer associated human cDNA sequence #321.
XX
KW   Colon cancer; T cell; tumour protein; C34S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00065543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PS   Claim 1; SEQ ID NO 321; 87pp; English.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov

Db
```
Db         481 ATGGCTGGGCTCTATGGAGAACAACGACGTGTCCAAGGCGAGGGGGTTGGGTTCCGTCGTCGT 540
```

Qy
```
Qy        1403 ACACGAAGTTGCTGGAGA 1421
               ||||||||||||||||||
Db         541 ACACGAAGTTGCTGGAGA 559
```

SQ   Sequence 539 BP; 191 A; 106 C; 132 G; 105 T; 0 U; 5 Other;

Query Match   24.8%;   Score 520.8; DB 11; Length 539;
Best Local Similarity 98.7%;   Pred. No. 2:1e-80;
Matches 522; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   205 GCCGGCCGAGCCCATGTCCCGGGGGGCCCCCAGGAGGTGAACCGGCTCACGGAGAGCACCTAC 264
          |||||||||| || ||||||||||||||||||||||||||||||| |||||||||||||||||
Db   11 GCCGGCCGCANCCNNTGTCCCCGGGGGCCCCGAGGAGGTGAACANGGTCACGGANAGCACCTAC 70

Qy   265 CGGAANTGTTATGGAACAGTTCAATCCTGGGCGTGCGGAAATTTAATAAACCTGGGGAAAAAT 324
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 NGGAATGTTATGGAACAGTTCAATCCTGGGCGTGCGGAAATTTAATAAACCTGGGGAAAAAT 130

Qy   325 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGGATGGAGTGGGCC 384
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 TATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGAAAAGCCTACTACGGATGGAGTGGGCC 190

Qy   385 AAGATCGGTGAGATTGCCACTGGGGTCCCCGGTGTCAACTGAACAGTGCCGTCATA 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 AAGATCGGTGAGATTGCCACTGGGGTCCCCGGTGTCAACTGAACTGGGACAATGTCCTCATA 250

Qy   445 GAGATTTCAAGTACCCACAAGAAACTCAACGGAGAGTCTTCGATGAAAATTTTAAAAAATTC 504
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 GAGATTTCAAGTACCCACAAGAAACTCAACGGAGAGTCTTCGATGAAAATTTTAAAAAATTC 310

Qy   505 CACAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGTGAAAATATATGAAC 564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 CACAAAGAGATTATCCATGAGCTGGAGAAGAGATAGAACTTGACGTGAAAATATATGAAC 370

Qy   565 GCAACTCTAAAAAGATACCAAAACACAAGAATAAAATTAGAGTCTTTGGAGAAATCC 624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 GCAACTCTAAAAAGATACCAAAACACAAGAATAAAATTAGAGTCTTTGGAGAAATCC 430

Qy   625 CAAGCTGAGTTGAGCAAGATTCAGRAGGAAAAGCCAAGGAAGCCACTCAAATAT 684
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 CAAGCTGAGTTGAGCAAGATTCAGAGGAAAAAGCCAAGGAAGCCACTCAAATAT 490

Qy   685 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTCCGTCAGAGTG 733
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 GAACACAAAAGAAATTGAGTATGTGGAGACCGTTACTTCTCGTCAGAGTG 539

Search completed: March 20, 2006, 13:54:24
Job time : 861 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2006 Bioacceleration Ltd.

GenCore version 5.1.7

Run on: March 20, 2006, 13:39:51 ; Search time 8382 Seconds
(without alignments)
11699.562 Million cell updates/sec

Title: US-10-801-292-1
Perfect score: 2096
Sequence: 1 ctcctctgctcctcgaagaa.........aaaaaaaaaaaaaaaaaaaa 2096

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 2339541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	1475.8	70.4	1479	11	DQ051599	DQ051599 Homo sapi
2	1223.4	58.4	3245	4	AK004918	AK004918 Mus muscu
3	1189.2	56.7	2214	4	AK008203	AK008203 Mus muscu
4	950	45.3	1479	11	DQ051600	DQ051600 Pan trogl
5	852.2	40.7	909	5	BI153780	BI153780 AGENCOURT
6	830	39.6	848	5	BU184787	BU184787 AGENCOURT
7	811.4	38.7	895	5	BQ888698	BQ888698 AGENCOURT
8	803.4	38.3	814	6	CA450094	CA450094 UI-CF-FN0
9	794.8	37.9	877	5	BU184917	BU184917 UI-CF-FN0
10	792.2	37.8	916	5	BQ688252	BQ688252 AGENCOURT
11	775.2	37.0	812	5	BI333892	BI333892 AGENCOURT
12	768.4	36.7	779	5	BM981661	BM981661 UI-CF-EN1
13	762.6	36.4	943	5	BU506695	BU506695 AGENCOURT
14	760.4	36.3	996	5	BI335025	BI335025 AGENCOURT
15	757.2	36.1	933	2	BG761459	BG761459 AGENCOURT
16	751.8	35.9	880	3	BQ233322	BQ233322 AGENCOURT
17	739	35.3	739	7	CN392816	CN392816 17006002
18	732.8	35.0	763	5	BU689777	BU689777 UI-CF-FN0
19	724.2	34.6	876	5	BU176351	BU176351 AGENCOURT
20	722.8	34.5	756	8	DR005158	DR005158 TC113330
21	718.8	34.3	747	2	BI255527	BI255527 60297567O
22	717	34.2	717	3	BI335902	BI335902 TmdI19 Hu
23	713.4	34.0	720	2	BF732309	BF732309 naeo9a07.
24	707.8	33.8	725	2	BG682313	BG682313 60262963I
25	696.8	33.2	919	2	BG674550	BG674550 60220003B
26	688.2	32.8	712	5	BU633552	BU633552 UI-H-FL1-
27	686.4	32.7	924	5	BQ922694	BQ922694 AGENCOURT
28	676	32.3	676	3	BM177746	BM177746 K-EST0084
29	665.8	31.8	847	2	BG034422	BG034422 60230282A
30	663.4	31.7	690	1	AW473168	AW473168 xq12f05.x
31	661.8	31.6	690	1	AI924372	AI924372 wn60b08.x
32	647.6	30.9	662	2	BM968354	BM968354 UI-CF-DUI
33	647.4	30.9	989	2	BG033827	BG033827 60301207
34	638.2	30.4	688	7	AW374090	AW374090 PM4-BT054
35	637	30.4	648	6	CV027819	CV027819 6171 Full
36	612.2	29.2	680	2	BE722245	BE722245 601568088
37	607	29.0	914	2	BG335429	BG335429 60240408T
38	604.8	28.9	890	5	BU181833	BU181833 AGENCOURT
39	597.4	28.5	789	3	BI560834	BI560834 60325419A
40	595	28.3	902	5	BY708155	BY708155 BY708155
41	593.4	28.3	1009	5	BI332789	BI332789 60257962S
42	585.4	27.9	588	2	BG766055	BG766055 60273843A
43	583.8	27.9	863	5	BU501073	BU501073 AGENCOURT
44	581	27.7	581	3	BM852920	BM852920 K-EST0134
45	578.6	27.6	931	5	BQ960401	BQ960401 AGENCOURT

ALIGNMENTS

RESULT 1
DQ051599

LOCUS   DQ051599   1479 bp   DNA   linear   GSS 02-JUN-2005
DEFINITION   Homo sapiens LOC55971 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION   DQ051599
VERSION   DQ051599.1 GI:66904813
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Homo.
REFERENCE   1 (bases 1 to 1479)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
              Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
              White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE   A Scan for Positively Selected Genes in the Genomes of Humans and
              Chimpanzees
  JOURNAL   (er) PLoS Biol..3 (6), E170 (2005)
  PUBMED   15869325
REFERENCE   2 (bases 1 to 1479)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
              Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
              White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
              them based on alignment. Translation starts at the beginning of
              alignment.
FEATURES   Location/Qualifiers
     source   1..1479
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="?"
     gene   <1..>1479
                     /gene="LOC55971"
                     /locus_tag="HC19618"
ORIGIN

Query Match   70.4%; Score 1475.8; DB 11; Length 1479;
  Best Local Similarity   99.9%; Pred: No. 1e-290;
  Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    274 ATGGAACAGTTCAATCCTGGGCTGCGGGAAATTTAATAAAACCTGGGCAAAGATTATGAGAA 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAACAGTTCAATCCTGGGCTGCGGGAAATTTAATAAAACCTGGGCAAAGATTATGAGAA 60

Qy    334 GCTGTAAAGCCTATGAATCCTGGGCAGGAAAAGCCTACTACTAGGATGAGTGGCCAAGATCGGT 393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCTGTAAAGCCTATGAATCCTGGGCAGGAAAAGCCTACTACTAGGATGAGTGGCCAAGATCGGT 120

Qy    394 GAGAATTGCCACTCGGGTCCCCGGTGTCAACTGAACTGGGACAGTGTCCTCATAGAGATTTCA 453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGAATTGCCACTCGGGTCCCCGGTGTCAACTGAACTGGGACAGTGTCCTCATAGAGATTTCA 180

Qy    454 AGTACCCACAAGAAACTCAACAGAGAGTCTTGATGAAAATTTAAAAAATTCCACAAAGAG 513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGTACCCACAAGAAACTCAACAGAGAGTCTTGATGAAAATTTAAAAAATTCCACAAAGAG 240

Qy    514 ATTATCCCATGAGCTGGAGAAGAGATGAACTTGACGTGAAATATATGATTTGTGGTTGATAAG 573
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATTATCCATGAGCTGGAGAAGAGATGAACTTGACGTGAAATATATGATTTGTGGTTGATAAG 300

Qy    574 AAAAGATACCAAACAGAACAAGAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAAAGATACCAAACAGAACAAGAGAATAAATTAGAGTCTTTGGAGAAATCCCAAGCTGAG 360

Qy    634 TTGAAGAAGAATCAGAAGGAAAAGCCAAGGAAAGCCACTCAAATATGAACACAAA 693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTGAAGAAGAATCAGAAGGAAAAGCCAAGGAAAGCCACTCAAATATGAACACAAA 420

Qy    694 GAAATTGAGTATGTGGGAGACCGTTACTTCGTGCGTCAGAGAAAATTCATTGCA 753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAATTGAGTATGTGGGAGACCGTTACTTCGTGCGTCAGAGAAAATTCATTGCA 480

Qy    754 GATGGTTGCAAAGAGGCTCGCTTCGAAGAGAAGAGGCGCGCCTTCGCTTTCTGGTTGATAAG 813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATGGTTGCAAAGAGGCTCGCTTCGAAGAGAAGAGGCGCGCCTTCGCTTTCTGGTTGATAAG 540

Qy    814 CACTGTGGCTTTGCAAACACAATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACTGTGGCTTTGCAAACACAATACATTATTATCACTTACAGTCTGCAGAACTACTGAAT 600

Qy    874 TCCAAGCTGCCTCGTCGGTGGCAGGAGACCTGTTGTTGATGTGCCAGAGAAAATC 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCCAAGCTGCCTCGTCGGTGGCAGGAGACCTGTTGTTGATGTGCCAGAGAAAATC 660

Qy    934 ATGAATATGATCGAAGAAATAAAGAACCCCGCAGCCTCTACCCCCGTCTCTGGAACTCCTCAG 993
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATGAATATGATCGAAGAAATAAAGACCCCGCAGCCTCTACCCCCGTCTCTGGAACTCCTCAG 720

Qy    994 GCTTCACCCATGATCGAGAGAGCAAATGGTGGTTAGGAAAGATTACGACACACCTTTCTAAA 1053
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTTCACCCATGATCGAGAGAGCAAATGGTGGTTAGGAAAGATTACGACACACCTTTCTAAA 780

Qy   1054 TGCTCACCCAAAGATGCCCCCCGCCGGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 1113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TGCTCACCCAAAGATGCCCCCCGCCGGCTCCTTCAGGCAGAGCATATACCAGTCCCTTGATCGAT 840

Qy   1114 ATGTTTAATAACCCAGCCACGGCTCGCCCCGAATTCACACAAGGGTAAATAATTCAACAGGT 1173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATGTTTAATAACCCAGCCACGGCTCGCCCCGAATTCACACAAGGGTAAATAATTCAACAGGT 900

Qy   1174 ACTTCCGAAGATCCCAGTTTACAGCGGATCCAGTTTCGGTTGCAACGGGACTGAACATGATG 1233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACTTCCGAAGATCCCAGTTTACAGCGGATCCAGTTTCGGTTGCAACGGGACTGAACATGATG 960

Qy   1234 AAGAAGCAGAAAGTGAAGAGACCATCTTCCCGCCACACTGCGTCCAACAAAGACCTTACTC 1293
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGAAGCAGAAAGTGAAGAGACCATCTTCCCGCCACACTGCGTCCAACAAAGACCTTACTC 1020

Qy   1294 AGCTTTGCACAGGGAGAGTGCATCACGCTGCTCATCCCGAGGGACGAAGAAGGACTGGGCTC 1353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCTTTGCACAGGGAGAGTGCATCACGCTGCTCATCCCGAGGGACGAAGAAGGACTGGGCTC 1080

Qy   1354 TATGGAGAACAGACGACGGTGTCCAAGGCAGGGCAGGGGGTTGGGTTCCGGTTCGTACGGTACGGAAGTTG 1413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TATGGAGAACAGACGACGGTGTCCAAGGCAGGGCAGGGGGTTGGGTTCCGGTTCGTACGGTACGGAAGTTG 1140

Qy   1414 CTGGAGAAAATGAGACAGAAGCAGTGACCGTGCCCAAGGCCAAGCCCCACACCAGTGAGA 1473
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTGGAGAAAATGAGACAGAAGCAGTGACCGTGCCCAAGGCCAAGCCCCACACCAGTGAGA 1200

Qy   1474 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCACGTGTTCATCCCCACCCACTAC 1533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCATCAGCACCGTGAACTTGTCTGAGAATAGCACGTGTTCATCCCCACCCACTAC 1260

Qy   1534 TTGGAATGCTTGTCCATGGGGGGCACGACTGCCGACAGGAGACAGATTCGGCGGACGGACGACA 1593
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTGGAATGCTTGTCCATGGGGGGCACGACTGCCGACAGGAGACAGATTCGGCGGACGGACGACA 1320

Qy   1594 TCCACCTTTAAGGCCCCAGGGGTCCAAGCCCGGCGCGGGGTCCTTAACGATGCCAACGGG 1653
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCCACCTTTAAGGCCCCAGGGGTCCAAGCCCGGCGCGGGGTCCTTAACGATGCCAACGGG 1380

Qy   1654 ACTGCAAAAGCCGCCCTTTTCTCAGCCGGAGAAAACCCCTTGCCACTGTGAAACTCCGCCCG 1713
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1381 ACTGCAAAAGCCGCCCTTTTCTCAGCCGGAGAAAACCCCTTGCCACTGTGAAACTCCGCCG 1440

Qy   1714 ACTGTGACGAAATGAATCGGCTCGGCACCACCATCATTCGATGA 1752
          |||||||||||||||||||||||||||||||||||||||||||
Db   1441 ACTGTGACGAAATGAATCGGCTCGGCACCACCATCATTCGATGA 1479
```

RESULT 2
AK004918

LOCUS   AK004918   3245 bp   mRNA   linear   HTC 03-APR-2004
DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched
            library, clone:1300006M19 product:INSULIN RECEPTOR TYROSINE KINASE
            SUBSTRATE homolog [Homo sapiens], full insert sequence.
ACCESSION   AK004918
VERSION   AK004918.2 GI:26338407
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tsehiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5

AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 3245)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT   On Dec 10, 2002 this sequence version replaced gi:12836466.
Please visit our web site (http://genome.gsc.riken.jp/) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGAGAGAGCGGCCGCCGACTCGAGTTTTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence[5'
GAGAGAGAGAGCCATCAAAGACGCTCAATTAATTAAATTAAAACCCCCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.
FEATURES   Location/Qualifiers
source   1..3245
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1300006M19"
/clone="1300006M19"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
CDS   273..1817
/note="unnamed protein product; INSULIN RECEPTOR TYROSINE
KINASE SUBSTRATE homolog [Homo sapiens] (SPTR|Q9UHR4,
evidence: FASTY, 90.3%ID, 99.2%length, match=1206)
putative"
/codon_start=1
/protein_id="BAB23669.1"
/db_xref="GI:12836467"
/translation="MSRGPEEVNRLTENTYRNVMEQFNPGLRNLINLGKNYEKAVNAM
ILAGKAYYDGVAKIGEIATGSPVSTELGHVLIEISSTHKKLNETLDENFKKHKDIIH
ELEKKTELDVKRWNATLKRYQAEHRNKLDSLEKSQAELKKIRRKSQGGRNALKYEHKE
IEYVETVTSRQSEIQKFIADGCKKALLEEKRFCFLVDKHCSPASHIHYYHMQSAELL
NSKLPRWQETCCDATKVPEKIMNMIEIEHTPISTPVSGTPQPSPMIERSKMIGKDYDT
LSKYSPKMPAPSVKAYTSPLIDMFNNPATAAQSSEKTNNSTANTGEDPSLQRSVSVA
TGLNMMKKQKYKTIFPHTAGNNKTLLSFAQGDVLTLLIPEEKDGWLYGEHDTTKARGW
FPSSYTKLLEENEAMSYVPTPSPAPVRSISTVDLTEKSSVVIFPPDYLECLSMGATSDK
KAGAPKVPSASTFRAPVSRPDATSTSPSDANGTAKPFPLSGENPPATVKLRPTVTNDR
SAPIIR"
polyA_signal   3227..3232
/note="putative"

ORIGIN   polyA_site   3245
        /note="putative"

Query Match   Best Local Similarity   58.4%;   Score 1223.4;   DB 4;   Length 3245;
Matches 1600;   Conservative   77.9%;   Pred. No. 3.9e-239;
0;   Mismatches 421;   Indels 34;   Gaps 9;

```
Qy      24  CAGGGCGGGGGCTGCCCGGAAGTTTTGACATTTTCGCGACGCGAGACGGCGGCCGCGGACACTCTC   83
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84  CAGACCGAAGCGCGCCGCAAAAGTTTTGACATTTTGGCAGCGAGACGACGGCGACTGACTCTC   143

Qy      84  GGGCCGACCGGCTGCGCGGGCGGGGGCGGCGGGACCCCCCAGAGCGCCCCTTAGTCGCGGCCGCCCGGCCCTC   143
            ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     144  CCGCTGACAGCTGTGTGGGCGGCCGCCGACTCTCAAGGACCACTCAGT----CTCAAGTCCTA   199

Qy     144  CCGGCTGCCCGGAGTCCGGCGGCCCACGAGGCCCCGGCGGGCCCCGGGTCCTCCCGCGGCCTTGCTCGGCC   203
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db     200  GGGCGCTCCGGCCGCGCCGGCACAGGGCCCACGAGGCCCGGCGGGCCCGGGTCCTCCCGCGGCGTTGCTCGGCC   259

Qy     204  GGCGGCCGCGGCAGCCCATTCTCCCGCGGGCGGGGCGGGGGGTGGAACCCCGGCTCACGGACACACTA   263
            |||||||||||||||||||||||||||||||||||||||||||||
Db     260  GGAGGCTGCAGCCATGTCTCGGGCGCCGGAGGTGAACCGGCTCACGGAGAACACGTA   319

Qy     264  CCGGAATGTTATGGAACAGTTCAATCCTGGGCTGCGAAATTTAATAAACCTGGGGGAAAAA   323
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db     320  CCGGAATGTGATGGAACAGTTCAACCCTGGACTACGGAAAACTTAATAAACCTAGGGAAAA   379

Qy     324  TTATGAGAAAGCTGTAAACGCTATGATCCTGGCAGGGAAAAGCTACTACGATGGGGTGGC   383
            ||||||||||||||||||||||||||||||||||||||||||
Db     380  TTATGAGAAAGCGGTGAATGCTATGATCCTGGCAGGGAAAGCCTACTACGACGGCGTGGC   439

Qy     384  CAAGATCGGTGAGATTGCCACTGGGTCCCCGGTGTCAAACTGAACGGACAGTGCATCTCTCAT   443
            |||||||||||||||||||||||||||||||||||||||
Db     440  CAAGATTGGTGAGATTGCCACTGGGTCTCCTGTGTCAAACAGAGCTAGGCCAATGCCATCAT   499

Qy     444  AGAGATTTCAAGTACCCACCAAAGAAAACTGGAGAGTGTCTTGATGAAAATTTTAAAAAT   503
            |||||||||||||||||||||||||||||||||||||||||
Db     500  AGAGATTTCCAGCACCCACAAGAAAACTCAATGAAGACTCTTGATGAAAATTTCAAAAATT   559

Qy     504  CCACAAAAGAGATTATCCATGAGCTGGAGAAGAAGACTTGACGTGAAATATATGAA   563
            ||||||||||||||||||||||||||||||||||||||||||||
Db     560  CCATAAGGATATTATCCATGAGCTGGAGAAGAAGACAGAACTTGATGTGAAAGTATATGAA   619

Qy     564  CGCAACTTCTACAAAAAATCAAACAAAACAGAGATAAATTAGAGTTCTTTGGAGAAATC   623
            ||||||||||||||||||||||||||||||||||||||||||
Db     620  CGCTACTCTTAAAAAGATATCAAGCAGAGCACAGGAATAAATTAGATTCTTTGGAGAAATC   679

Qy     624  CCAAGCTGAGTTGAGAAGATTCAGAAGAAAAAGCCAAGGAAGCCCACTCAAATA   683
            ||||||||||||||||||||||||||||||||||||||||||
Db     680  TCAGCGCGGAGCTGAAGAAGATCAGAAGAAAAGTCAAGGTGGACACGCACTCAAATA   739

Qy     684  TGAACACAAAGAAAATTGAGTATGTGGAGACCGGTTACTTCTCGTCAGAGTGAAATCCAGAA   743
            ||||||||||||||||||||||||||||||||||||||||||
Db     740  CGAGCACAAAGAAAATTGAGTATGTCGGAGACGGTGAACGTCTGCCCAGAGTGGATCCAGAA   799

Qy     744  ATTCAТTGCAGATGGTTGCAAAGGCGCTTCTGCTTCTGCTTCT   803
            ||||||||||||||||||||||||||||||||||||||
Db     800  GTTCAТTGCCGATGGCTGCAAAGGAAGGCGCTTTTGCTTTCT   859

Qy     804  GGTTGATAAGCACTGTGGCTTGCAAACCCACATACTATTCACTTACAGTCTGCAGA   863
            ||||||||||||||||||||||||||||||||||||||
Db     860  GGTGGACAAGCACTGTGCAGTCTTTGCCAGTCACTATCATCAGTCTCCAGA   919

Qy     864  ACTACTGAATTCCAAGCTGCCTCGGTGGCGAGGAGGACCTGTGTTGATGCCATCAAAGTGCC   923
            ||||||||||||||||||||||||||||||||||||||
Db     920  GTTACTTAATTCCAAGCTGCCCCGGTGGCAGGAAACCTGTTGTGACGCCACAAAAGTCC   979

Qy     924  AGAGAAAATCATGAATATGATTGAGAAGAAATAAAGACCCCAGCCTTACCCCGTGTCTGG   983
            ||||||||||||||||||||||||||||||||||||||
Db     980  AGAGAAAATCATGAACATGATTAGAAGAAGAGATAAGCGCCCAATCTCCACCCCCGTGTCTGG   1039

Qy     984  AACTCCTCAGGCTTCACCCCATGATCGAGAGAAGCAATGTGGTTAGGAAAAAGATTACCGACAC   1043
```

Db 2106 AAATAGCCAGAATAA 2120

```
Qy  1040 GACGCGCGCAGCCATCGCCCATGCCATCGATCGAGGAGGAAGCCAGATGATTGGGAAAGATTATGACAC 
              |||   ||  |||         |||      ||      ||  |||||||||    ||||
Db  1099 CCTTTCTAAATGCTCACCAAAAGATGACCCCCCCCCCCTTCCTTCAGGCAGGAGCATATACCAGTCC 1103

Qy  1044 CCTTTCTAAATACTCACCAAAAGATGACCCCCCCCCCCTCCTTCCTTCAGGCAGGAGCATATACCAGTCC 
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  1100 CCTTTCTAAATACTCACCAAAAGATGACCCCCCCCCCCTTCCTTCAGGCAGGAGCATATACCAGTCC 1159

Qy  1104 CTTGATCGATATGTTTAATAACCCAGCCACGGGCTGCCCGGGAATTCACAAAAGGGTAAATAA 
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1160 TTTGATTGACATGTTCAATAAAACCCAGCTCCAGCGTCACAAGAGTTCAGAAAAAAAACAAACA 1163

Qy  1160 TTTGATTGACATGTTCAATAAAACCCAGCTCCAGCGTGCACAAGAGTTCAGAAAAAAAACAAACA 
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219

Qy  1164 TTCAACAG----GTACTTCCGAAGATCCCAGTTTACAGCGCATCAGTTTGCAAACGGG 
              ||||||      ||||||||||||||||||||||||||||||||||||||||
Db  1220 CTCCACAGCAAACACTGGGGAAGATCCCAGTCTACAGCGGGTTTCTGTTGCAACTGG 1220

Qy  1220 CTCCACAGCAAACACTGGGGAAGATCCCAGTCTACAGCGGGTTTCTGTTGCAACTGG 
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279

Qy  1221 ACTGAACATGATGAAGAAAGGCAGAAAGTGAAGAGCTAGAAGACCATTCTTTTCCCGCACACACTGCGGGGCTCCAA 
              ||||  ||      ||       |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1280 GCTGAAACATGATGAAGAAAGGCAGAAAGTGAAGAGCTAGAAGACCATTCTTTTCCCGCACACACTGCGGGGCTCCAA 1280

Qy  1280 GCTGAAACATGATGAAGAAAGGCAGAAAGTGAAGAGCTAGAAGACCATTCTTTTCCCGCACACACTGCGGGGCTCCAA 
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339

Qy  1281 CAAGACCTTACTTACTTTGCACAGGGGAGGAGTGCATCAGGAGTGTCATCCCCGAGGAGAA 
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1340 TAAGACCTACTGAGCTTTGCCCAGGGGAGCGTCCTCACACTTCTTCATCCCTGAGGGAGAA 1340

Qy  1340 TAAGACCTACTGAGCTTTGCCCAGGGGAGCGTCCTCACACTTCTTCATCCCTGAGGGAGAA 
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399

Qy  1341 GGATGGCTGGCCTCTATGGAGAACACGAGCTGTCCAAGGGGGAGGGGGTTGGTTCCCGTCGTC 
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1400 GGACGGGCTGGCCTCTACGGGGAGCATGAACACCACCAAAGCCGAGGGGCTGGTTCCCATCGTC 1400

Qy  1400 GGACGGGCTGGCCTCTACGGGGAGCATGAACACCACCAAAGCCGAGGGGCTGGTTCCCATCGTC 
              ||||||||||||||||||||||||||||||||||||||||||||||
Db  1459

Qy  1401 GTACGAGAGTTGCTGGAGAAGAAATGAGACAGAGAAAATGAGACACGATGGGTTTCTGTTCACACCAAGCCC 
              ||||||      ||||||||  |||||||||||||||  |||||||||||||||||||||
Db  1460 ACTTCCGACTACTTGGAGTGCCTTGTCTCCATGGGAGACCACTCAGACAAAGAGAAAGGGGGGGC 1460

Qy  1460 ATATACAAAGTTGTGTGGGAGAAAAT-----GAAGCAATGTCTGTGTCCCTTACACCAAGCCC 
              ||||||                     |||||||||||||||||||||||||||||
Db  1461 ACACACAGTGAGGAGCATCAGCACAGCACCGTGAACTTGTTCTGAGAATAGCAGTTGTTGTCATCCC 1513

Qy  1461 CACACACAGTGAGGAGCATCAGCACAGCACCGTGAACTTGTTCTGAGAATAGCAGTTGTTGTCATCCC 
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1520

Qy  1514 TGCACCCGTGAGGAAGGCATCAGCGTGGGGGAGCATGAGCATGACTGAGAAGGAGCAGGCCTTGTGTCATCC 
              ||||||||||||||||||||||||||||||||||||||||||||||
Db  1521 CCCACCCGACTACTTGGAATGCTTGTCATGGGGGGCCGACAGCAGACGAGGAGCAGAGATTC 1573

Qy  1521 CCCACCCGACTACTTGGAATGCTTGTCATGGGGGGCCGACAGCAGACGAGGAGCAGAGATTC 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  1580

Qy  1574 ACTTCCGACTACTTGGAGTGCCTTGTCTCCATGGGGGGAGACCACTCAGACAAAGAGAAAGGGGGGGC 
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1581 GGCCAGG-------ACGACATCCACACTTTAAGGCCCCAGTCCAAGCCCCA-----GAC 1633

Qy  1581 GGCCAGG-------ACGACATCCACACTTTAAGGCCCCAGTCCAAGCCCCA-----GAC 
              ||||||                                           ||||
Db  1628

Qy  1629 CGGGGCTCCTAACGATGCCCAAGGGACTGCCAAAGCCGGGGCTCTTTTCTCAGGCGGGGAAAACCC 
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1634 TCCCAAGGTGCCTTCTGCTTCTTACCTTCAGCTGCCCAGTGTCCAGCACGAGCCCACATC 1693

Qy  1634 TCCCAAGGTGCCTTCTGCTTCTTACCTTCAGCTGCCCAGTGTCCAGCACGAGCCCACATC 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  1688

Qy  1689 CTTTGCCACTGTGAAAACTTCCGCCCGCACGAGAGCCGGGGCACCGCACCCATCATTCG 
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1694 CACCTCTCCCAGCGGATGCCACAGCGGGGACCCGGAAAAGCCGCCCAGTGACCACATCATTGAA 1748

Qy  1694 CACCTCTCCCAGCGGATGCCACAGCGGGGACCCGGAAAAGCCGCCCAGTGACCACATCATTGAA 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  1753

Qy  1749 ATG-AGAGACAGCACAGGACTCTCCCGGGCCCTCTCCCGTTCTCCCGGGGATGGG 
              |||  ||||||||||||||||||||||||||||||||||||||||||||
Db  1754 CTTTGCCACGGTGAAAACTCCGGCCGTGACCAGTGACAAAATGACCGATCCATCCG 1813

Qy  1754 CTTTGCCACGGTGAAAACTCCGGCCGTGACCAGTGACAAAATGACCGATCCATCCG 
              ||||||||||||||||||||||||||||||||||||||||
Db  1807

Qy  1808 GCGGCATCCTGTTCTCGCCACGTGCTGCTGCGACCGGTCGGGAAGCTTCAGTGTGGGAGAGGCCTAACTCTA 
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1814 ATGAAGATCGACAGGAGCAGCAGTGCCAGTGCCAGTTCTCACCCACAGGCGGATAG 1873

Qy  1814 ATGAAGATCGACAGGAGCAGCAGTGCCAGTGCCAGTTCTCACCCACAGGCGGATAG 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  1867

Qy  1868 ATGTCGCTCGCTTCTTAGCTAAAATCATGCCTCGTTCACGTGTTGACGTTGGGGTTTGGGGTTGACAAGTTTC 
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1874 GTATA-CATTGCTCGGACACAGCGGCACTGTCAAAAGATTGCAGAGGTGTCCTGTGATGTCAATGAA 1932

Qy  1874 GTATA-CATTGCTCGGACACAGCGGCACTGTCAAAAGATTGCAGAGGTGTCCTGTGATGTCAATGAA 
              ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  1927

Qy  1928 TGCCTTTAAGGATAATAGAGTAATAGTCTCTAATGACCAGCTCAGATATTAAAATATTTTCT 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  1933 GCCCCACTTGAGCAAATCAGAATCAATGTCGTTTTCAGTGTGGGTTGTGACAAGTTTG 1992

Qy  1933 GCCCCACTTGAGCAAATCAGAATCAATGTCGTTTTCAGTGTGGGTTGTGACAAGTTTG 
              ||||||||||||||||||||||||||||||||||||||||||
Db  1987

Qy  1988 TCCTATTCTGTTCAAGAGAAACAGTAAACTTGGGTTTCAATCTTTAAAAAAAAAAAAAAAAA 
              ||||||||||||||||||||||||||||||||||||||||||
Db  1993 TGTCTTTAGTGAATAAATGAATAAATCATTGAATATAATTTTCT 2049

Qy  1993 TGTCTTTAGTGAATAAATGAATAAATCATTGAATATAATTTTCT 
              ||||||||||||||||||||||||||||||||||||||||||||
Db  2047

Qy  2048 AAAAAAAAAAAATAA 2062 
              |||  ||   |||
Db  2050 T----TTCTATTCAAAAAGCAGTAAATTTGGCTCAAATTTTTAAATTTTTTATTTTTAAGG 2105
```

RESULT 3

LOCUS   AK008203   2214 bp   mRNA   linear   HTC 03-APR-2004
DEFINITION   Mus musculus adult male small intestine cDNA, RIKEN full-length
             enriched library, clone:2010012D17 product:INSULIN RECEPTOR
             TYROSINE KINASE SUBSTRATE homolog [Homo sapiens], full insert
             sequence.
ACCESSION   AK008203
VERSION   AK008203.1 GI:12842246
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
             Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 2214)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
             Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
             Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
             Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
             Kasukawa,T., Kato,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
             Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
             Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
             Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
             Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
             Tejima,Y., Toya,T., Yamamura,Y., Yasunishi,A., Yoshida,K.,
             Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
             URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,

```
COMMENT     Fax:81-45-503-9216)
            Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Sciences Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAAGAGATCCAAGAGGATCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 5.0 and subtraction to
            Rot = 20.0. Second strand cDNA was prepared with the primer adapter
            of sequence[5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCC 3'], cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES             Location/Qualifiers
     source          1..2214
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2010012D17"
                     /db_xref="taxon:10090"
                     /clone="2010012D17"
                     /sex="male"
                     /tissue_type="small intestine"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    177..1717
                     /note="INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE homolog
                     [Homo sapiens] (SPTR|Q9UHR4, evidence: FASTY, 90.3%ID,
                     99.2%length, match=1206)
                     putative"
     polyA_signal    2192..2197
                     /note="putative"
     polyA_site      2214
                     /note="putative"
ORIGIN

Query Match     56.7%;  Score 1189.2;  DB 4;  Length 2214;
Best Local Similarity   77.8%;  Pred. No. 3.6e-232;
Matches 1586;  Conservative  0;  Mismatches 413;  Indels  39;  Gaps  11;

Qy      41 AAGTTTTGACATTTTCGCAGCGGAGACGGCGCGGGAGCACGTGCGGGC 100
           ||||||||||||||||   ||  || || || || ||| ||||||||||
Db       5 AAGTTTTGACATTTTGGGCAGCGAGCGGCCGCACTGAGTCTCTCCCGAC 64

Qy     101 GGGCGGCCGACCCTCCAGAGCCCCTTAGTCGCGGAGCGCGGAGAGTCCG 160
           ||||||||  |||    ||||  ||  ||| |   ||| ||||||||||
Db      65 GGGCGGCCGACTCTCAAGGACACACTCAGT---CTCAAGTCCTAGGCG 120

Qy     161 GCGCGGCCACGAGGGCCCCAGCGCGCGGCCGCGGCGCTTGCTCGGCCGCG 220
           ||||  |  |||||||||||  |||  ||  || || ||    ||||||
Db     121 GCGCCGCCAGGCCCCGTCTCCCCGCCGGCTTGCTCGGCGGCCCATGT 180

Qy     221 CCCGGGGGGCCCGAGGGAGGTGAACCGGCTCACGGAGCACCTCAGCAGGAAC 280
           ||||||   ||||||| | |||||| ||   |||||||||   ||| ||||
Db     181 CTCGGCGGGCCCCGGAGGAGGTTGAACCGGGAATGCTGATGGGAAC 240

Qy     281 AGTTCAATCCTGGGCTGCAAAATTTAATTAAAACCTGGGGAAAATTATGAGAAAAGCTGTAA 340
           | ||||||||  ||||   | |||| |||| ||  |    |||   ||||||| ||||||
Db     241 AATTCAAACCCTGGACTACGGAAATTAATAAACCTACTACGACGGCGTGA 300

Qy     341 ACGCTATGATCCTTCTGGGCCAGGAGAAAGCCTACTACGATGGACTGGCCAAGATCGGCTGAGATTG 400
           |||||||  ||||  ||| ||||||||||  || | |   ||||| ||||||| |||||||
Db     301 ATGCTATGATCCCTGGCAGGAAAAAGCCTACTACGACGGCGGTGGGCCA---TGGGTGAAATG 356

Qy     401 CCACTGGGTCCCCCGTGTCAACTGGACATGTCCTCATAGAGAATTTCAAGTACCC 460
           ||||| |||    ||| ||| ||| || || | | |    || |||||||  ||
Db     357 CCACTGGGTCTCCTCGTGTCAACAGAGCACGCTGTCTCATCCCGACTACTTGG 416

Qy     461 ACAAGAAACTCAACGAGAGTCTTGATGAAAAATTTTTAAAAATTCCACAAAGAGATTATCC 520
           |||| ||||| | ||||| | |||| || |   ||| | ||||||||||| ||||||  |
Db     417 ACAAGAAACTCAATGAGAGCTCTTGATGATGAAAAATTTCCATAAAGGATTATCC 476

Qy     521 ATGAGCTGGAGAAGAAGAGATAGAACTTGACGTGAAATATATGAACGAAGA 580
Db     477 ATGAGCTGGAGAAGAGAGAGTAGAGAGAACTTGATGTGAAACGGCTTAGT... 536

Qy     581 ACCAAACAGACCAACAAGAATAAAATTAGAAGTCTTTGGGAGAAATCCCAAGCTGAGTTGAGA 640
Db     537 ATCAAGCAGAGCAGAGCAGGAATAAATTAGATTCTTTTTGGGAGAAATTCAGGGCAGCTGAAGA 596

Qy     641 AGATCAGAAGAAAAGCCCAAGACGCCCAAAACGACCACTCAAAAATTGAAGACAAAAGAAATTG 700
Db     597 AGATCAGAAGAAAAGTCAAGGTGGACGAAAAACCACTCAAAAATACGAGCACAAAGAAAATTG 656

Qy     701 AGTATCGTGGGAGACCGGTTACTTCTCGTCGAGAGTTGAGAGTCAGGAAATTCATTGCAGATCGTT 760
Db     657 AGTATCGTGGGAGACGGGTGACGGTCTCCGCCAGATGGAGTGCATTGCCGCT 716

Qy     761 GCAAAGAGGCTCTGCTTGAAGAGGAAGGCGGCTTCTTCGGTTTCGGGTTTGATAAGCACTGTG 820
Db     717 GCAAAGAAGCCCCTGCTTGAGGAAAAGGCGCTTCTTTCTGCTTTCGGGTGGGACAAGCACTGCA 776

Qy     821 GCTTTTTCCAAAACCAACATACATTTATTATCACTTACAGTCTGCAGAACTACTGAATCCAAAGC 880
Db     777 GCTTTTGCCAGTTCAACATACTACTACTTACATGCAGTCTGCAGAGGTTACTTAATTCCAAAGC 836

Qy     881 TGCCTCTGGGTTGGCCAGGACCCTGTTGTTGATGCCATCAAAGTGCCAGGAAAAATCATGAATA 940
Db     837 TGCCCCGGTTGCCAGGAAACCTGTTGTGAATGCCACAAAAAGTCCCAGAGAGAAAATCATGAACA 896

Qy     941 TGATCGAAGAAATAAAGACACCCGAGGCCTCTACCCCGTGTCTGGGAACTTCCTCAGGCTTCAC 1000
Db     897 TGATAGAAGAGATAAAAGACGCGCCCATCTCACACCCCGTGTCCGGCGCAGCCATCGC 956

Qy    1001 CCATGATCGGAGAGAAGCAAATGTGGTTGGGAGGAAAGGATTACGGACACACCCTTCTAAATGCTCAC 1060
Db     957 CCATGATCGGAGAGAAGCAAGATGATTGGGAAGAGATTATGACACCCTTTCTAAATACTCAC 1016

Qy    1061 CAAAGAATGCCCCCCCGTCTCCTTCAGGCAGAGCATATACCAGTGCCCTTGATCGATATGTTTTA 1120
Db    1017 CAAAGATGGCCCCCCTGCTTCCTTCAGTCAAAAGCCTACACCAGTCCTTTGACATTGTTCA 1076

Qy    1121 ATAACCCAGCCACGCCCTGCCCCGGAATTCACAGAGGGGGTAAATAATTCAACAG---GTACTT 1177
Db    1077 ATAACCCAGCTACAGCTGCACAGAGATTTACATCCTCAAGAGAAGCAACTCCACAGCAAACACTG 1136

Qy    1178 CCGAAGATCCCAGTTTACAGGGGATCAGTTCGCGTGCAACATGATGAAGA 1237
Db    1137 GGGAAGAGATCCCAGTCTACAGCTCTACAGCGGATCGGTTTCAACATGATGAAGA 1196

Qy    1238 AGCAGAAAGTGAAGACCAATCTTTCCCCGCCACACTGCGGGCTCCAACAGACCTTACTCAGCT 1297
Db    1197 AGCAGAAGGTGAAGACCAATCTTTCCTCAACAGGCTGGCAACATAGACCTTACTGAGCT 1256

Qy    1298 TTGCACAGGGAGATGTCATCAGCTGCCTCCATCCCCGAGGAGAAAGTGTTATGGAAC 1357
Db    1257 TTGCCCAGGGGGGACAGTTCCTCAGCACTGTCTCCGAGGGATGAACGGCTTCTTACG 1316

Qy    1358 GAGAACAGCAGGTGTCCAAAGCGCGAGGGGTTGGGTTCCCGTGGCTGGCTGG 1417
Db    1317 GGGAGCCATGACACCCAACACCAAAGCGCGAGGGCTTGGTTCCCATGCTGTCGTGGCTGG 1376

Qy    1418 AAGAAAAATGAGGACAGAAGCAGTGACCGTGCCCAGCCCAGCCCCACCAGTGAGAGAGCA 1477
Db    1377 AAGAAAAT---------GAAGCAATGTCTTGTCCCTACCAACAAGCCCCTGACCGTGAGAGCA 1430

Qy    1478 TCAGCACCGTTGAACTTGTTCTTGAGAATAGCAGTGTGTTGTCATCCCCACCACTGCTTGG 1537
Db    1431 TCAGCACCGGTGACCTGACCTGAGAGCAGCACCGTTGTCATCCCCACCACTTGG 1490

Qy    1538 AATGCTTGTCCATGGGGGCAGCTGCCGCACAGGAGAGCAGATTCGGGCCAGG-------ACGA 1591
           ||||                                                          |
Db    1531 ...
```

```
RESULT 4
DQ051600
LOCUS,        DQ051600        1479 bp    DNA    linear    GSS 02-JUN-2005
DEFINITION    Pan troglodytes LOC55971 gene, VIRTUAL TRANSCRIPT, partial
              sequence, genomic survey sequence.
ACCESSION     DQ051600
VERSION       DQ051600.1   GI:66904814
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
  ORGANISM    Pan troglodytes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Pan.
REFERENCE     1 (bases 1 to 1479)
  AUTHORS     Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
              Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
              White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE       A Scan for Positively Selected Genes in the Genomes of Humans and
              Chimpanzees
  JOURNAL     PLoS Biol. 3 (6), E170 (2005)
  PUBMED      15869325
REFERENCE     2 (bases 1 to 1479)
  AUTHORS     Nielsen,R., Bustamante,C., Clark,A.G. Glanowski,S., Sackton,T.B.,
              Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
              White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE       Direct Submission
  JOURNAL     Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT       This sequence was made by sequencing genomic exons and ordering
              them based on alignment. Translation starts at the beginning of
              alignment.
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>1479
                     /gene="LOC55971"
```

ORIGIN   /locus_tag="HC19618"

```
Query Match
Best Local Similarity   45.3%;    Score 950;  DB 11;  Length 1479;
Matches  953;  Conservative  64.4%;  Pred. No. 2.6e-183;  Mismatches 526;  Indels  0;  Gaps  0;
```

```
Qy    274 ATGGAACAGTTCAATCCTGGGCTGCCAAATTTATAAAACCTTGGGGAAAATTATGAGAAA 333
          ||||||||||   |||||| |||||||||||||||     |||||||||||||||||||||
Db      1 ATGGAACANTTCAATCCTTGGGCTGCCAAATTNANNNNNNCCTACTACGATGGGAAAATTATGAGAAA 60

Qy    334 GCTGTAAAACGCTTATGATCCTGGCAGGAAAAGCCTACTACGATGGGACGAAAAATTATGAGGGT 393
          || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||| |||
Db     61 GCTGTAAACGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120

Qy    394 GAGAATTGCCACTGGGGTCCCCCGTGTCAACTGAACTGGGACATGTCCTCATAGAGATTTCA 453
          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGACATGTCCTCATAGAGATTTCA
Db    121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
```

```
Db     961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020

Qy    1294 AGCTTTGCAAGGGACGAGATGTCATCACGCTGCTCATCCCGAGGAGGAAGGATGGCTCGGCTC 1353
Db    1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080

Qy    1354 TATGAGAACACGACCTGTCCAAGGGCCGAGGGGTTGGTTCCCGTCGTCGTACACGGAAGTTG 1413
Db    1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140

Qy    1414 CTGGAAGAAAATGAGACAGAGAAGCAGTGACCGTGCCCACGCGCCCCACACCAGTGAGA 1473
Db    1141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCAGTGAGA 1200

Qy    1474 AGCAATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTGTTGTCATCCCCCCACCCGACTAC 1533
Db    1201 AGCAATCAGCACCGTGAACTTGTCTGAGAATAGCAGTGTGTTGTCATCCCCCACCCGACTAC 1260

Qy    1534 TTGGAATGCTTGTCCATGGGGGGCCAGCTGTGCCACACAGAGAGAGATTCGGCCAGGACGACA 1593
Db    1261 TTGGAATGCTTGTCCATGGGGGGGCAGCTGCCCGACAGGAGAGAGGATTCGGCCAGGACGACA 1320

Qy    1594 TCCACCTTTAAGGCCCCAGGTCCAAGCCCGAGACCCGGCGGCCTCCTAAGGATGCCAACGGG 1653
Db    1321 TCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1380

Qy    1654 ACTGCAAAGCCGCCTTTCTCAGGGGAGAAAACCCCCTTGCCCACTGTGAAAACTCCGCCCG 1713
Db    1381 NNNNNNNNNNNNNNNNNNNNCCGGAGAAAACCCCTTGCCCACTGTGAAAACTCCGCCCG 1440

Qy    1714 ACTGTGACGAAATGATCGGCTCGGCACCCACCCATCATTCGATGA 1752
Db    1441 ACTGTGACGAAATGATCGGCTCGGCACCCACCCATCATTCGATGA 1479
```

```
ORIGIN        Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.75 kb. Library constructed by Life
              Technologies."

Query Match   40.7%; Score 852.2; DB 5; Length 909;
Best Local Similarity 96.8%; Pred. No. 2.3e-163;
Matches 880; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
```

```
Qy     960 CCCAGGCCTTCACCCCGTTGTCTGGAACTTCCTCAGGCTTCACCCATGATCGAGGAAGCAA 1019
Db       1 CCACGCGCTTCCCGCCCCACGCGGTCCGCCCCACAGGCGCCCATGATCGAGGAAGCAA 60

Qy    1020 TGTGGTTAGGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCGTCC 1079
Db      61 TGTGGTTAGGGAAAGATTACGACACCCTTTCTAAATGCTCACCAAAGATGCCCCCGTCC 120

Qy    1080 TTCAGGCGAGAGCAATATACCACTCCCTTGATCGATATGTTTAATAACCCAGCCACGGTGCC 1139
Db     121 TTCAGGCGAGAGCAATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCACGGTGCC 180

Qy    1140 CCCGACACATCTCGGGGCTCCAACAAGACCTTACTCAGCTTTCCACCAGGGAGATCCCAGTTTACACAC 1199
Db     181 CCCGAATTCACAAAGGGTAAAATAATTCAACAGGTAGTTACTTCCGAGAGATCCCAGTTTACAGCG 240

Qy    1200 ATCAGTTTCGGTTGCAACGGGACACTGAACATGATGAAGAAGACCACTCTT 1259
Db     241 ATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGACCACTCTT 300

Qy    1260 CCCGCACACTCGGGGCTCCAACAAGACCTTACTCAGCTTTCCACCAGGGAGATGTCATCAC 1319
Db     301 CCCGCACACTCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGAGATGTCATCAC 360

Qy    1320 GCTGCTCATCCCCGAGGAGGAAGGATGGCTCGGCTCTATGAGGAGAACACGACCGTGTCCAAGGC 1379
Db     361 GCTGCTCATCCCCGAGGAGGAAGGATGGCTCTATGAGGAGAACACGACGGTGTCCAAGGC 420

Qy    1380 GAGGGGGTTGGGTTCCCGTCGTCGTACACGGAAGTTGCTGGAAGAAAATGAGACAGAGAAGCAGT 1439
Db     421 GAGGGGGTTGGGTTCCCGTCGTCGTACACGGAAGTTGCTGGAAGAAAATGAGACAGAGAAGCAGT 480

Qy    1440 GACCCTGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGA 1499
Db     481 GACCCTGTGCCCACGCCCAAGCCCCACACCAGTGAGAAGCATCAGCACCGTGAACTTGTCTGA 540

Qy    1500 GAATAGCAGTGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGGCAGC 1559
Db     541 GAATAGCAGTGTGTTGTCATCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGGGCAGC 600

Qy    1560 TGCCGGACAGGAGAGGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCCAGGCGTCCAA 1619
Db     601 TGCCGGACAGGAGAGGAGCAGATTCGGCCAGGACGACATCCACCTTTAAGGCCCCAGGCGTCCAA 660

Qy    1620 GCCCGAGACCCGGCGGCCTCCTAAGGATGCCAACGGGACTGCAAAGCCGCCCTTTCTCAGGGG 1679
Db     661 GCCCGAGACCCGGCGGCCTCCTAAGGATGCCAACGGGACTGCAAAGCCGCCCTTTCTCAGGGG 720

Qy    1680 AGAAAACCCCTTTGCCCACTGTGAAACTCGTGACGAAATGATCGGCTCGGCGACC 1739
Db     721 AGAAAACCCCTTTGCCCACTGTGAAACTCCGCCCGACTGTGACGAAATGATCGGCTCAGCACC 780

Qy    1740 CATCATTCGATGAGGAGGACAGCCAAGGAGACTCTCCCGGAGCTTCCCGGTTCTCCCTTGCGG 1799
Db     781 CATCATTCGATGAGGAGGACAGCCAAGGAGACTCTCCCGGAGCTTCCCGGTTCTCCCTTGCGG 840

Qy    1800 AATGATGGGCGCATCCTGTCTGCCACGTGC-TGACGGGTCGGGAAGCTTCAGTGGATGGAGAGGC 1858
Db     841 AATGATGGGCGCGCCGTCTTGCCCGGTGCTTGAGGTCGGGGAACTTTCAGTGGTAGGAGAGGC 900

Qy    1859 CTAACTCTA 1867
Db     901 CTAACTCTA 909
```

```
LOCUS       BU153780                909 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7780106 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137888
            5', mRNA sequence.
ACCESSION   BU153780
VERSION     BU153780.1  GI:22667312
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/;
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13452 row: g column: 09
            High quality sequence stop: 732.
FEATURES             Location/Qualifiers
     source          1..909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6137888"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
```

RESULT 5
BU153780

```
RESULT 6
BU184787
LOCUS       BU184787                 848 bp    mRNA    linear    EST 04-SEP-2002
DEFINITION  AGENCOURT_7825409 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6093297
            5', mRNA sequence.
ACCESSION   BU184787
VERSION     BU184787.1  GI:22698771
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13362 row: e column: 10
            High quality sequence stop: 710.
FEATURES             Location/Qualifiers
     source          1..848
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6093297"
                     /tissue_type="epithelioid carcinoma"
                     /clone_lib="NIH_MGC_70"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.1 kb. Library constructed by Life
                     Technologies."
ORIGIN

Query Match             39.6%;  Score 830;  DB 5;  Length 848;
Best Local Similarity   99.8%;  Pred. No. 7.9e-159;
Matches  841;  Conservative    0;  Mismatches   1;  Indels   1;  Gaps   1;

Qy        953 TAAAGACCCCAGCCCTCTACCCCCGTGTCTCGAAACCCCCAGCTTCACCCATGATCGAGA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAAAGACCCCAGCCCTCTACCCCCGTGTCTCGAAACCCCCAGCTTCACCCATGATCGAGA 60

Qy       1013 GAAGCAAATGTGGTTAGGAAAGAATTACGACACACCCTTTCTAAATGCTCACCAAAGATGCCCCC 1072
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGCAAATGTGGTTAGGAAAGAATTACGACACACCCTTTCTAAATGCTCACCAAAGATGCCCC 120

Qy       1073 CCGGCTCCTTCAGGCAGGAGCATATACCCAGTCCCTTGATCGATATGTTTAATTAACCCAGCCA 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGCTCCTTCAGGCAGGCAATATACCAGTCCCTTGATCGATATGTTTAATTAACCCAGCCA 180

Qy       1133 CGGGCTCCCCCGAAATTCACAAAGGGTAAAATAATTCAACAGGTACTTCCGGAAGGAAGATGGG 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGCTCGCCCCGAAATTCACAAAGGGTAAAATAATTCAACAGGTACTTCCGGAAGGAAGATGGG 240

Qy       1193 TACAGGGGATCAGTTTCCGTTTGCAACGGGACTGAACAGCAGAAAGTGAAGA 1252
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACAGGGGATCAGTTTCCGTTTGCAACGGGACTGAACAGCAGAAAGTGAAGA 300

Qy       1253 CCATCTTCCCGCCACACTGCGCGGCCCTCAACACAGACCTTACTCACAGGGGAGATG 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCTTCCCGCCACACTGCGCGGCCCTCAACACAGACCTTACTCACAGGGGAGATG 360

Qy       1313 TCATCAGCGCTGCTCATCCCCGAGGAGGAAGGATGGCTGGCTCTATGGAGAACGACGTGT 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCATCAGCGCTGCTCATCCCCGAGGAGGAAGGATGGCTGGCTCTATGGAGAACGACGTGT 420

Qy       1373 CCAAGGCGGAGGGGTTGGTTCCCGTTGCTGTAGCAGCGAAGAAGTTGCTGGAAGAAAAATGAGACAG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAGGCGGAGGGGTTGGTTCCCGTTGCTGTAGCAGCGAAGAAGTTGCTGGAAGAAAAATGAGACAG 480

Qy       1433 AAGCAGTGACCGTGCCCACGCCCAAGCCCCACCACCAGTGAGAAGCATCAGCACCGTGAACT 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGCAGTGACCGTGCCCACGCCCAAGCCCCACCACCAGTGAGAAGCATCAGCACCGTGAACT 540

Qy       1493 TGTCTGAGAATAGCAGTGTTGTTCATCCCCCACCCGACTACTTGTCCATGG 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTCTGAGAATAGCAGTGTTGTTCATCCCCCACCCGACTACTTGTCCATGG 600

Qy       1553 GGGGCAGCTCCCGACAGCAGGAGAGCAGAGAGATTCCGGCCAGGACGACATCACCCTTTAAGGCCCAG 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGGCAGCTCCCGACAGCAGGAGAGCAGAGAGATTCCGGCCAGGACGACATCACCCTTTAAGGCCCAG 660

Qy       1613 CGTCCAAGCCCCGAGACCGCGCGGTCCTTACCGATTGCGAACGGGAGTCCAAAGCCCAGCTTTTC 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCAAGCCCCGAGACCGCGCGGTCCTTACCGATTGCCAAGCGGGAGTCCAAAGCCCAGCTTTTC 720

Qy       1673 TCACGCGGAGAAAAACCCCTTTGCCACTGTGAAAACTCCGCCCGACTGTGACGAAATGATCGGCT 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCACGCGGAGAAAA--CCCTTTGCCACTGTGANACTCCGCCCGACTGTGACGAAATGATCGGCT 779

Qy       1733 CGGCACCCATCATTCGATGAGAGGAGCAGCAGCCAAGAGACTCTCCCGGGGCCTCTCCCGGGGTTCTCC 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CGGCACCCATCATTCGATGAGAGGAGCAGCAGCCAAGAGACTCTCCCGGGGCCTCTCCCGGGGTTCTCC 839

Qy       1793 CTT 1795
              |||
Db        840 CTT 842
```

```
RESULT 7
BQ888698
LOCUS       BQ888698                 895 bp    mRNA    linear    EST 16-AUG-2002
DEFINITION  AGENCOURT_8074718 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6086238
            5', mRNA sequence.
ACCESSION   BQ888698
VERSION     BQ888698.1  GI:22280712
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2320 row: o column: 07
            High quality sequence stop: 705.
FEATURES             Location/Qualifiers
     source          1..895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6086238"
                     /tissue_type="ductal carcinoma, cell line"
                     /clone_lib="NIH_MGC_110"
                     /lab_host="DH10B (phage-resistant)"
```

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 38.7%; Score 811.4; DB 5; Length 895;
Best Local Similarity 98.2%; Pred. No. 5e-155;
Matches 863; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

```
Qy    309 AAACCTGGGGAAAAAATTATGAGAAAGCTGTAAACGGTATGATCCTGCTGCAGGAAAAGCCTA 368
          ||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      4 AACCTGGGGGAAAAATTATGAGAAAGCTGTAAACGGTATGATCCTGGCTGCAGGAAAAGCCTA 63

Qy    369 CTACGACTGGAGTGGCCAAGATCGGTGACATTGGCCACTGGCCTTGTCAACTGAACT 428
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 CTACGACTGGAGTGGCCAAGATCGGTGAGATTGGCCACTGGCCTTGTCAACTGAACT 123

Qy    429 GGGACACATGTCCTCATAGAGAGATTTCAGTACCACACAGAGAACTCAACGAGAGTCTTGATGA 488
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 GGGACACATGTCCTCATAGAGATTTCAGTGACCACACAAGAAACTCAACGAGAGTCTTGATGA 183

Qy    489 AAATTTTAAAAAATTCCACAAAGAGAGATTATCCATGAGCTGTGGAGAAGAGATAGAACTTGA 548
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 AAATTTTAAAAAATTCCACAAAGAGATTATCCATGAGCTGTGGAGAAGAGATAGAACTTGA 243

Qy    549 CGTGAAATATATGAACGGCAACTCTAAAAAAGAGATTACCCATGAGGTGGGAAGAGAATAAATTAGA 608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 CGTGAAATATATGAACGGCAACTCTAAAAAAGAGATTACCCATGAGGTGGGAAGAGAATAAATTAGA 303

Qy    609 GTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGAGGAAAAAGCCAAGGAAGCCG 668
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 GTCTTTGGAGAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGAGGAAAAAGCCAAGGAAGCCG 363

Qy    669 AAAACGGCACTCAAAATATGAACACAAAGAAAATTGAGTAGTGTGGAGACCCGTTACTTCTCGTCA 728
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 AAAACGGCACTCAAAATATGAACACAAAGAAAATTGAGTAGTGTGGAGACCCGTTACTTCTCGTCA 423

Qy    729 GAGTGAAATCCAGAAAATTCATTGCGACGATGGTTGTTGCAAAGAGGCTCTCGCTTGAAGAGAAGAG 788
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GAGTGAAATCCAGAAAATTCATTGCGACGATGGTTGTTGCAAAGAGGCTCTCGCTTGAAGAGAAGAG 483

Qy    789 GCGCTTCTCGCTTCTTCTGGTTGATAAGGACTGTGGCCTTTGCGAAAACCACATACATTATTATTATTCA 848
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 GCGCTTCTCGCTTCTTCGTTGATAAGCCACTGTGGCCTTTGCAAAGAGGCTTACTCATTATTATCA 543

Qy    849 CTTACAGTCTGCGAGAACTACTGAATTCCAAGCTGCCTCGGTGCCAGGAGACCTGTGTTGA 908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTTACAGTCTGCGAGAACTACTGAATTCCAAGCTGCCTCGGTGCCAGGAGACCTGTGTTGA 603

Qy    909 TGCCATCAAAGTGCCAGAGAAAATCATGAATCGAGAAAAATAAAGACCCCAGCCTC 968
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 TGCCATCAAAGTGCCAGAGAAAATCATGAATCGAGAAAAATAAAGACCCCAGCCTC 663

Qy    969 TACCCCCGGTGTCTGGAACTCCTCAGCCTCACCCATGATCGAGGAAGCAATGTGGTTAG 1028
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TACCCCCGGTGTCTGGAACTCCTCAGCCTCACCCATGATCGAGGAAGCAATGTGGTTAG 723

Qy   1029 GAAAGATTACGGACACCTTCTAAATGCTCACCAAAGATG-CCCCCGCTCCTTCAGGCA 1087
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    724 GAAAGATTACGGACACCTTTCTAAATGCTCACCAAAGATGCCCCCCGCTCCTTCAGGCA 783

Qy   1088 GAGCATATACCAGTCCCTTGATGCGATATGTTTAAT-AACCCAGCCACGGCTG-CCCCGAA 1145
          |||||||||||||||||||||||||||||||||||| |||||||||||||| |||||||
Db    784 GAGCATATACCAGTCCCTTGATCGATATGTTTAATAAACCCACCGACGGCTGCCCCCGAA 843

Qy   1146 TTCACAAAAGGGTAAA--TAATTCAACAAGGTACTTCCGAA 1182
          ||||| ||| |||||     |||||||||||||||||||||
Db    844 ATCACCAAGGGGAAAAATAAATTCACAAGGTACTTTCGAA 882
```

RESULT 8

```
LOCUS       CA450094                814 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-CF-FN0-afb-o-02-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
            UI-CF-FN0-afb-o-02-0-UI 3', mRNA sequence.
ACCESSION   CA450094
VERSION     CA450094.1  GI:24814514
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 814)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M.J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
FEATURES             Location/Qualifiers
     source          1..814
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-FN0-afb-o-02-0-UI"
                     /tissue_type="Human Lung Epithelial cells"
                     /lab_host="DH10B (life Technologies)"
                     /clone_lib="UI-CF-FN0"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-FN0 is a subtracted cDNA library derived from two
                     normalized Human lung epithelial cell libraries (EN1 and
                     DU1) The library was subtracted according to according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. For additional information, contact:
                     bento-soares@uiowa.edu
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-FN0
                     TAG_SEQ=CTGCTCAGGT"
            POLYA=Yes.
ORIGIN
```

Query Match 38.3%; Score 803.4; DB 6; Length 814;
Best Local Similarity 99.1%; Pred. No. 2.1e-153;
Matches 807; Conservative 0; Mismatches 2; Indels 7; Gaps 0;

```
Qy   1239 GCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGGGGTCCAACAAGACCTTACTCAGCTT 1298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 GCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGGCTCCAACAAGACCTTACTCAGCTN 755

Qy   1299 TGCACAGGGAGGATGTCATCACGCTGCTCATCCCCGAGGAGGAAGGACTGTGGCCTCTATGG 1358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 TGCACAGGGAGGATGTCATCACGCTGCTCATCCCCGAGGAGGAAGGACTGTGGCCTCTATGG 695

Qy   1359 AGAACACGACGTGTCCAAGGCGAGCGGAGGGGTTGGTTCCGTCGTCGTTCGTACACGAAGTTGCTGGA 1418
```

```
RESULT 9
LOCUS       BU184917     877 bp    mRNA    linear    EST  04-SEP-2002
DEFINITION  AGENCOURT_7825544 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6093473
            5', mRNA sequence.
ACCESSION   BU184917
VERSION     BU184917.1  GI:22698901
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13362  row: 1  column: 18
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6093473"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.1 kb. Library constructed by Life
                     Technologies."
ORIGIN
Query Match  37.9%;  Score 794.8;  DB 5;  Length 877;
Best Local Similarity 97.8%;  Pred. No. 1.2e-151;
Matches 859; Conservative  0; Mismatches 12; Indels 7; Gaps 5;
```

```
Qy      1118 TTAATAACCCAGCCACGGCGCTCGCCCCGGAATTCACAAGGGTAAAATAATTCAACAGGTACTT 1177
             ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         1 TTAATAACCCAGCCACGGCGCTCGCCCCGGAATTCACAAAGGGTAAAATAATTCAACAGGTACTT 60

Qy      1178 CCGAAGATCCCAGTTTTACAGCCATCACTTTTCGGTTGCAAACGGGACTGAACATGATGAAGA 1237
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CCGAAGATCCCAGTTTTACAGCGATCAGTTTCCGGTTGCAAACGGGACTGAACATGATGAAGA 120

Qy      1238 AGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCGCTCCAACAAGACCTTACTCAGCT 1297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AGCAGAAAGTGAAGACCATCTTCCCGCACACTGCGGGCGCTCCAACAAGACCTTACTCAGCT 180

Qy      1298 TTCCACAGGGAGGATGTCATCACGCTCTCATCCCCGAGGAGGAAGGATGGCTGGCTCTATG 1357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TTGCCACAGGGAGGATGTCATCACGCTCTCATCCCCGAGGAGGAAGGATGGCTGGCTCTATG 240

Qy      1358 GAGAACACGACACTGTCCAAGGCCGAGGGGGTTGGGTTCCCGTCGTCGTACACGAAGTTGCTGG 1417
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GAGAACACGACGTGTCCAAGGCCGAGGGGGTTGGGTTCCCGTCGTCGTACACGAAGTTGCTGG 300

Qy      1418 AAGAAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCA 1477
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AAGAAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCA 360

Qy      1478 TCAGCACCCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCACCCGACTACTTGG 1537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TCAGCACCCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCACCCGACTACTTGG 420

Qy      1538 AATGCTTCCATGGGGGCCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCA 1597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AATGCTTGTCCATGGGGGCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCA 480

Qy      1598 CCTTTAAGGCCCCAGCGGTCCAAGCCCGAGACCCGGCGGTCCTCTAACGGATGCCAACCGGACTG 1657
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CCTTTAAGGCCCCAGCGGTCCAAGCCCGAGACCCGGCGGTCCTCTAACGGATGCCAACCGGACTG 540

Qy      1658 CAAAGCCGCGCTTTTCTCAGCGGGCGAGAAAAACCCCTTTGCCACTGTGAAAACTCCGCCCGACTG 1717
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CAAAGCCGCGCTTTTCTCAGCGGGCGAGAAAAACCCCTTTGCCACTGTGAAAACTCCGCCCGACTG 600

Qy      1718 TGACGAATGATCCGCTCGGCACCCATCATTCGATGATGAGAGGACAGCGCAAGGACTCTCCCGGG 1777
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TGACGAATGATCGGCTCGGCACCCATCATTCGATGATGAGAGGACAGCGCAAGGACTCTCCCGGG 660

Qy      1778 CCTCTCCCGGTTCTCCCTTGCGGAATGATGGGCGCCATCCGTGCTCCACGTGCTGACGGGTC 1837
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCTCTCCCGGTTCTCCCTTGCGGAATGATGGGCGCCATCCGTGCTCCACGTGCTGACGGGTC 720

Qy      1838 GGGGAAGCTTCAGTGGGAGAGGGGCCCTAACTCTAAATGTCGCGCCTTAAGCAAAATCATGTCTTCT 1897
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       720 GGGGAAGCTTCAGTGGGAGAGGGGCCCTAACTCTAAATGTCGCCCTGCTTAAGC-AATCATGCTTCT 778
```

```
Qy        34 TTTCAATCTTTACTGTAAAAAAAAAAAAA 2052
             ||||||||||||||| ||||||||||||
Db      2019 TTTTCAATCTTTTACTGTAAAAAAAAAAATAATAAAAAAAAA 1

Qy        94 GACTAGCTTCAGCGCCATTTAAAATATTTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGG 35
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1959 GACCAGCTCAGCCATTTAAAATATTTTCTTCCTATTCTGTTCAAGAAACAGTAAACTTGG 2018

Qy       154 TGTTTCACGTAGTTGGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTTAAT 95
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      1899 TGTTTCACGTAGTTGGGTTGACAAGTTTCTGCCTTTAAGATAAATGAGTAATAGTCTTAAT 1958

Qy       214 GGAAGCGTTCTGTGGAGAGGCCTAACTCTAAATGTCGCGCCTTAAGCAAAATCATGCTTCTC 155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1839 GGAAGCTTCAGTGGGAGAGGGGCCTAACTCTAAATGTCGCGCCTTAAGCAAAATCATGCTTCTC 1898

Qy       274 CTCTCCCGGTTCTCCCTTGCGGAATGATGGGCGCCATCCGTGCTCCACGTGCTGACGGGTCG 215
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1779 CTCTCCCGGTTCTCCCTTGCGGAATGATGGGCGCCATCCGTGCTCCACGTGCTGACGGGTCG 1838

Qy       334 GACGAATGATCGGCTCGGCACCCATCATTCGATGATGAGAGGACAGCGCAAGGACTCTCCCGGC 275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1719 GACGAATGATCCGCTCGGCACCCATCATTCGATGATGAGAGGACAGCGCAAGGACTCTCCCGGC 1778

Qy       394 AAAGCCGCGCTTTTCTCAGCGGAGAAAAACCCCTTTGCCACTGTGAAAACTCCGCCCGACTGT 335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1659 AAAGCCGCGCTTTTCTCAGCGGAGAAAAACCCCTTTGCCACTGTGAAAACTCCGCCCGACTGT 1718

Qy       454 CTTTAAGGCCCCAGCGGTCCAAGCCCGAGACCCGGCGGTCCTAACGGATGCCAACGGGACTGC 395
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1599 CTTTAAGGCCCCAGCGGTCCAAGCCCGAGACCCGGCGGTCCTAACGGATGCCAACGGGACTGC 1658

Qy       514 ATGCTTGTCCATGGGGGCCAGCTGCCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCAC 455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1539 ATGCTTGTCCATGGGGGCCAGCTGCCGACAGGAGAGCAGATTCGGCCAGGACGACATCCAC 1598

Qy       574 CAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCACCCGACTACTTGGA 515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1479 CAGCACCGTGAACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCACCCGACTACTTGGA 1538

Qy       634 AGAAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCAT 575
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1419 AGAAAAATGAGACAGAGAGCAGTGACCGTGCCCACGCCAAGCCCCACACCAGTGAGAAGCAT 1478

Qy       694 AGAACACGACCTGTCCAAGGCCGAGGGGGTTGGTTCCCGTCGTCGTACACGAAGTTGCTGGA 635
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      ...
```

```
RESULT 10
LOCUS       BQ688252
BQ688252
DEFINITION  AGENCOURT_8034808 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6207009
            5', mRNA sequence.
ACCESSION   BQ688252
VERSION     BQ688252.1  GI:21813568
KEYWORDS    EST.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 916)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     Unpublished (1999)
  JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2362 row: g column: 10
            High quality sequence stop: 673.

FEATURES             Location/Qualifiers
     source          1..916
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6207009"
                     /tissue_type="ductal carcinoma, cell line"
                     /clone_lib="NIH_MGC_110"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: Pancreas; Vector: pOTB7; Site 1: XhoI;
                     Site 2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC library."
```

ORIGIN

```
Qy          1898 CTGTTTCAC--GTAGTTGGGTTGACAAGTTT--CTGCCTTTAAGATAAATGAGT-AATAG 1952
                 |||||||||  ||||||||||||||||||||  |||||||||||||||| ||||| |||||
Db           779 CTGTTTCCCGTAGTTGGGTTGACCAGTTTCTGCCTTTTAAGATAAATGAGTAAATAG 838

Qy          1953 TCTTAATGACCAGCTCAGCCATTTAAAATATTTCTTCC 1990
                 |||||||||||||||||||||||||||||||||| ||
Db           839 TCTTAATGACCAGCTCAGCCATTTAAAATATTTCCTTC 876
```

Query Match

Best Local Similarity   37.8%;
Matches 794; Conservative   0; Mismatches   3; Indels   0; Gaps   0;
Score 792.2; DB 5; Length 916;
99.6%; Pred. No. 4.2e-151;

```
Qy    308 TAAACCTTGGGGAAAAATTATCAGAAAAGCTGTTAAACCCGTTATGATCCTCTGGCAGGAGAAAAAGCCT 367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TAAACCTTGGGGAAAAATTATGAGAAAAGCTGTTAAACCCGTTATGATCCTCTGGCAGGAGAAAAAGCCT 60

Qy    368 ACTACCAATGGGAGAGTGGGCCAAGATCGGTGGGTGAGGATTGGCCACTGGGTCCCCCGTGTCAACTGAAC 427
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTACCAATGGGAGAGTGGGCCAAGATCGGTGGGTGAGGATTGGCCACTGGGTCCCCCGTGTCAACTGAAC 120

Qy    428 TGGGACAGTGTCCTCATAGAGATTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGATG 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGGACAGTGTCCTCATAGAGATTCAAGTACCCACAAGAAAACTCAACGAGAGTCTTGATG 180
```

```
RESULT 11
LOCUS       BI333892
BI333892
DEFINITION  602999431F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141574 5',
            mRNA sequence.
ACCESSION   BI333892
VERSION     BI333892.1  GI:15018549
KEYWORDS    EST.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     Unpublished (1999)
  JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.B. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11348 row: n column: 07
            High quality sequence stop: 812.
```

```
Qy    488 AAAATTTTAAAAAATTCACAAAAGAGAGATTATCCATGAGCTGGAAGAAGAAGATAGAACTTG 547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAATTTTAAAAAATTCACAAAAGAGAGATTATCCATGAGCTGGAAGAAGAAGATAGAACTTG 240

Qy    548 ACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCAAAACAACACAAGAGATAAAATAG 607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACGTGAAAATATATGAACGCAACTCTAAAAAAGATACCAAAACAACACAAGAGATAAAATAG 300

Qy    608 AGTCTTTGGAGAAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCC 667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGTCTTTGGAGAAAATCCCAAGCTGAGTTGAAGAAGATCAGAAGGAAAAGCCAAGGAAGCC 360

Qy    668 GAAACGCACTCAAAATATGAACCAAGAAAGAAATTGAGTAGTGGAGAACCGTTACTTCGTC 727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAAACGCACTCAAAATATGAACCAAGAAAGAAATTGAGTAGTGGAGAACCGTTACTTCGTC 420

Qy    728 AGAGTGAAAATCCAGAAAATTCATTGCAGATGGTTGCAAAGACGGCTCTGCTTGAAGACGA 787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGAGTGAAAATCCAGAAAATTCATTGCAGATGGTTGCAAAGACGGCTCTGCTTGAAGACGA 480

Qy    788 GGGCCTTCGTCGCTTTCTGGGTTTGATAAGCACTGTGGTTGCAAAACCACATACATTATATC 847
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGGCGTTCGTCGCTTTCTGGGTTGATAAGCACTGTGGCTTTGCAAACCACATACAATTATATC 540

Qy    848 ACTTACAGGTCTGCAGGAACTACTGAATTCCAAGCTGCCTCGGTGGCGGAGGACCTGTGTTG 907
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ACTTACAGGTCTGCAGGAACTACTGAATTCCAAGCTGCCTCGGTGGCGGAGGACCTGTGTTG 600

Qy    908 ATGGCGTACAAAGTGCCAGAAGAAAAATCAGATGAATATGATCGAAGAAAATAAAAGAAGACCCCAGCCT 967
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATGCCATCAAAGTGCCAGAAGAAAAATCAGATGAATATCGAAGAAAAAATAAAAGAACCCCAGCCT 660

Qy    968 CTACCCCCGTGTCTGGAACTCCCTCAGGCGCTTCACCCATGATCGAGGAGAAGCAGAATGTGGGTTA 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTACCCCCGTGTCTGGAACTCCCTCAGGCTTCACCCATGATCGAGGAGAAGCAGAATGTGGGTTA 720

Qy   1028 GGAAAGATTACGGACACCCCTTTCTAAATGCTCACCAAAGATGCCCCCCGTCCTTCAGGCCA 1087
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGAAAGATTACGGACACCCCTTTCTAAATGCTCACCAAAGATGCCCCCCGTCCTTCAGGCCA 780

Qy   1088 GAGGCATATACCAAGTCCC 1104
          |||||||||||||||||
Db    781 GAGGCATATACCAAGTCC 797
```

FEATURES
```
source          Location/Qualifiers
                1..812
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5141574"
                /tissue_type="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH MGC 12"
                /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.4 kb. Library prepared by Life
                Technologies."
```

ORIGIN

```
Query Match
Best Local Similarity    37.0%;       Score 775.2; DB 3; Length 812;
Matches 799; Conservative  0;  Pred. No. 1.2e-147;
                              Mismatches  3;  Indels  2;  Gaps  2;
```

```
Qy   709 GAGACCGTTACTTCTTCCTCCAGAGTGAAATCCAGAAAATTCATTGCAGATGGGTTGCAAAGAG 768
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAGACCGTTACTTCTTCCTCCAGAGTGAAATCCAGAAAATTCATTGCAGATGGGTTGCAAAGAG 60

Qy   769 GCTCTGCTTGAAGAGAAGAGAGGCGCCTTCTTCGCTTTCTTTCGGTTGATAAGCACTGTGGGCTTTGCA 828
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCTCTGCTTGAAGAGAAGAGAGGCGCCTTCTTCGCTTTCTTTCGGTTGATAAGCACTGTGGGCTTTGCA 120

Qy   829 AACCACATACATTATTATCACTTACAGTCTCGCGAGAACTTACTGAATTCCAAGTCTGCCTCGG 888
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCACATACATTATTATCACTTACAGTCTCGCGAGAACTTACTGAATTCCAAGTCTGCCTCGG 180

Qy   889 TGGCAGGAGACCTGTGTTGATGGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATCGAA 948
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGGCAGGAGACCTGTGTTGATGGCCATCAAAGTGCCAGAGAAAATCATGAATAATGATCGAA 240

Qy   949 GAAATAAAGACCCCAGCCTCTACCCCGTGTTCTTGGACTCCTCCAGGCTTCACCCATGATC 1008
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAATAAAGACCCCAGCCTCTACCCCGTGTTCTTGGACTCCTCCAGGCTTCACCCATGATC 300

Qy  1009 GAGAGAAGCAAATGTGGTTAGGAAAAGAATTACGACACCCTTTCTTAAATGCTCACCAAAGATG 1068
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGAGAAGCAAATGTGGTTAGGAAAAGAATTACGACACCCTTTCTTAAATGCTCACCAAAGATG 360

Qy  1069 CCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCCTTGATCGATATATGTTTAATTAACCCA 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCCCCCGCTCCTTCAGGCAGAGCATATACCAGTCCCCTTGATCGATATATGTTTAATTAACCCA 420

Qy  1129 GCCACGGCGCTGCGCCCGGAATTCACACAAGAGGGTAAAATAATTAAATTCAACACAGGTAGTTACTTCTGAAGATCC 1188
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCACGGCGCTGCGCCCGGAATTCACACAAGAGGGTAAAATAATTAAATTCAACACAGGTAGTTACTTCTGAAGATCC 480

Qy  1189 AGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTG 1248
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGTTTACAGCGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTG 540

Qy  1249 AAGACCATCTTCCCGCACACTGCGGGCGGCTCCAACAAAGACCTTACTCTCAGCTTTGCACAGGA 1308
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAGACCATCTTCCCGCACACTGCGGGCGGCTCCAACAAAGACCTTACTCTCAGCTTTGCACAGGA 600

Qy  1309 GATGTCATCACGCTGCTCATCCCCGAGGAAGAAGGATGCGGCTGGCTTCCTTATGGAGAACACGAC 1368
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GATGTCATCACGCTGCTCATCCCCGAGGAAGAAGGATGCGGCTGGCTTCCTTATGGAGAACACGAC 660

Qy  1369 GTGTCCAAGGCGAGGGGTTGGTTCCCGTGCGTGGTACCGAAGTTGCTGGAAGAAAATGAG 1428
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GTGTCCAAGGCGAGGGGTTGGTTCCCGTGCGTGGTACCGAAGTTGCTGGAAGAAAATGAG 720

Qy  1429 ACAGAAGCAGTGACGTGCCCACGCCAAGCCCCACCAGTGAGAAGCAT-CAGCACCGT 1488
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAGAAGCAGTGACCGGTG-CCACGCCAAGACCCACACCAGTGAGAAGCATCAGCACCGT 779

Qy  1488 GAACTTGTCTGAGAATAGCAGTGT 1511
         |||||||||||||||||||||||
```

Db   780 GAACTTGTCTGAGAATAGCAGGGT 803

```
LOCUS       BM981661          779 bp    mRNA    linear   EST 21-FEB-2003
DEFINITION  UI-CF-EN1-adi-m-10-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-adi-m-10-0-UI 3', mRNA sequence.
ACCESSION   BM981661
VERSION     BM981661.1  GI:19604378
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M.J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            The following repetitive elements were found in this cDNA
            sequence: 1-21, >AT_rich#Low_complexity
            Seq primer: M13 FORWARD
            POLYA=Yes.
```

FEATURES
```
source          Location/Qualifiers
                1..779
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-CF-EN1-adi-m-10-0-UI"
                /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                Cells"
                /dev_stage="Adult"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /clone_lib="UI-CF-EN1"
                /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                modified polylinker; Site_1: EcoR I; Site_2: Not I;
                UI-CF-EN1 is a normalized cDNA library containing the
                following tissue(s): Primary Lung Cystic Fibrosis
                Epithelial Cells. The library was constructed according to
                Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                1996. First strand cDNA synthesis was primed with an
                oligo-dT primer containing a Not I site. Double stranded
                cDNA was ligated to an EcoR I adaptor, digested with Not
                I, and cloned directionally into pT7T3-Pac vector. The
                oligonucleotide used to prime the synthesis of the
                first-strand cDNA contains a library tag sequence that is
                located between the Not I site and the (dT)18 tail. The
                sequence tag for this library is CTGTCACGGT.
                6hr to LPS 24h
                TAG LIB=UI-CF-EN1
                TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                TAG_SEQ=CTGTCACGGT"
```

ORIGIN

```
Query Match
Best Local Similarity    36.7%;       Score 768.4; DB 3; Length 779;
                         99.1%;       Pred. No. 3e-146;
```

RESULT 13

BU500695

BU500695

```
LOCUS       BU500695    943 bp    mRNA    linear    EST   12-SEP-2002
DEFINITION  AGENCOURT_7860835 NIH_MGC_64 Homo sapiens cDNA clone IMAGE:6109543
            5', mRNA sequence.
ACCESSION   BU500695
VERSION     BU500695.1  GI:22802021
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     Unpublished (1999)
  JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: NIMH/LOG
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2355 row: j column: 08
            High quality sequence start: 43
            High quality sequence stop: 453.
FEATURES             Location/Qualifiers
     source          1..943
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6109543"
                     /clone_lib="NIH_MGC_64"
                     /note="Vector: pOTB7a; Site 1: CeuI; Site 2: SceI; This
                     library is a size selection of NIH_MGC_35, from 3.0-4.5
                     kb. Size selection done at the National Institute of
                     Mental Health, NIH. Note: this is a NIH_MGC Library."
                     /lab_host="DH10B (phage-resistant)"
ORIGIN
```

Query Match   36.4%;   Best Local Similarity   95.3%;   Score 762.6;   DB 5;   Length 943;
Matches 865;   Conservative 0;   Mismatches 32;   Indels 11;   Gaps 8;   Pred. No. 4.7e-145;

Matches 772;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

```
Qy   1269 TGCGGGCTTCCAACAAGACCTTACTCACGCTTTGCACAGGAGGAAGATGTCATCACGCTGCTCAT 1328
          |||||||||||||||||||||||||| |||| ||||||||||| ||||||||||||||||||||
Db    779 TGCGGGCTTCCAACAAGACCTTACTCAGCTTGCNACAGGAGGAGATGTCATCACGCTGCTCAT  720

Qy   1329 CCCGGAGGAAGGATGGCTGGCTTCTATGGAGGAGAACACGACGGTGTTCCAAGGCGAGGGGTTG 1388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    719 CCCGGAGGAAGGATGGCTGGCTTCTATGGAGGAGAACACGACGGTGTTCCAAGGCGAGGGGTTG  660

Qy   1389 GTTCCCGTCGTGTACACGGTAGCAGGTTGCTGGAAGAAGAAAATGAGAGACGAGAAGCAGTGACC 1448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    659 GTTCCCGTCGTGTACACGGTAGCAGGTTGCTGGAAGAAAATGAGACACGAGAAGCAGTGACCGTGCC  600

Qy   1449 CACGCCCAAGCCCCACACCAGTGAGAAGTTGTCTGAGAATAGCAG 1508
          ||||||||||||||||||||||||||||||||||||||||||||
Db    599 CACGCCCAAGCCCCACACCAGTGAACTTGTCTGAGAATAGCAG  540

Qy   1509 TGTTGTCATCCCCCACCCGACTACTTGGAATGCTTGTCCATGGGGCCAGCTGCCGACAG 1568
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 TGTTGTCATCCCCCACCCGACTACTTGGAATGCTTATCCATGGGGCCAGCTGCCGACAG  480

Qy   1569 GAGAGCAGAGATTCGGGCCAGGAGACAGCACATCCACCCTTTAAGGCCCCAGGGTCCAAGCCCGAGAC 1628
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GAGAGCAGAGATTCGGGCCAGGAGACCGACATCCACCCTTTAAGGCCCCAGGGTCCAAGCCCGAGAC  420

Qy   1629 CCGGGGCTCCTAACGATGCCAAGGGGACCCCTTTTCTCAGCGGAGAAAACCC 1688
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CGCGGGCTCCTAACGATGCCAAGGGGACCCCCTTTTCTCAGCGGAGAAAACCC  360

Qy   1689 CTTTGCCACTGTGAAACTTCCCGCCGACTGTGACGGAATGATGGCTCGGCAACCCATCAGTTCG 1748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 CTTTGCCACTGTGAAACTTCCCGCCGACTGTGACGAATGATGGCTCGGCACCCATCAGTTCG  300

Qy   1749 ATGAGAGGACAGCCAAGGAGACTCTCCCGGGCTCCTCTCCCGGGTTCTCCCTTGCGGAATGATGGG 1808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 ATGAGAGGACAGCCAAGGACTCTCCCGGGCTCCCTCCCGGGTTCTCCCTTGCGGAATGATGGG  240

Qy   1809 CGCATCCTGTCGCACGTGCTGACGGGTCGGGGAAGCTTCAGTGACGAGAGGCCCTAACTCTAA 1868
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CGCATCCTGTCGCCAAGGACGTGCTGACGGGTCGGGGAAGCTTCAGTGACGAGAGGCCCTAACTCATC  180

Qy   1869 TGTCGCGCTGCTTAAGCAAATCATGCTTCTCTCGTTTCACGTAGTTGGTTGACAAGTTTCT 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 TGTCGCGCTGCTTAAGCAAATCATGCTTCTCTCGTTTCACGTAGTTGGTTGACAAGTTTCT  120

Qy   1929 GCCTTTAAGATAAATGAGTAATAGTCTAATGACCAGCTCAGGCCATTAAAATATTTTCTT 1988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 GCCTTTAAGATAAATGAGTAATAGTCTAATGACCACTCACCCATTAAAATATTTTCTT   60

Qy   1989 CCTATTCTGTTCAAGAAGAACAGTAAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAA 2047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59 CCTATTCTGTTCAAGAAACAGTAAAACTTGGTTTCAATCTTTAAAAAAAAAAAAAAAAAAA    1
```

```
Qy   1151 AAAGGGTAAATATATTCAACAGGT-ACTTCCGAGAGATCCCCAGTTTACACGGATCAGTTTCG 1209
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     17 AACGGGGTAAATAATTCAACAGGTAACTTCTGAGAGATCCCAGTTTACACGGATCAGTTTCG   76

Qy   1210 GTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTCATCTTCCCGCCACACT 1269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGACCATCTTCCCGCCACACT  136

Qy   1270 GCGGGGCTCCAACAAGACCTTACTCACGCTTTGCACAGGAGGAGATGTCATCACGCTGCTCATC 1329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 GCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGGGGAGAGTCATCACGCTGCTCATC  196

Qy   1330 CCCGGAGGAAGGATGGCTGGCTTCTATGGAGGAGAACACGACGGTGTTCCAAGGCGAGGGGTTGG 1389
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 CCCGGAGGAAGGATGGCTGGCTTCTATGGAGGAGAACACGACGGTGTTCCAAGGCGAGGGGTTGG  256

Qy   1390 TTCCCGTCGTGTACACGGAGTGGTTGCTGGAAGAAGAAAATGAGAGACGAGAAGCAGTGACCGTGCCC 1449
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 TTCCCGTCGTGTACACGGAGTTGCTGGAAGAAGAAAATGAGAGACGAGAAGCAGTGACCGTGCCC  316

Qy   1450 ACGCCCAAGCCCCACACCAGTGAGAAGAAGCATCAGCACACCGTGAACTTGTCTGAGAATAGCAGT 1509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 ACGCCCAAGCCCCACACCAGTGAGAAACATCAGCACACCGTGAACTTGTCTGAAAATAGCAGT  376

Qy   1510 GTTGTCA-TCCCCCACCCGACTACTTGGAATGCTTGT---CCATGGGGCCAGCTGCCGA 1565
          ||||||| |||||||||||||||||||||||||||||    ||||||||||||||||||
Db    377 GTTGTCATTCCCCCACCCGACTACTTTGGAATGCTTGTCCATGGGGCCAGCTGCCGA  436

Qy   1566 CAGGGAGGAGCAGCAGATTCGGCCCAGGACGACGACA-TCCACCTTTAAGGCCCCCAGGGTCCAAGCCCG 1624
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    437 CAGGGAGGAACAGATTCGGCCCAGGACGACGACATCCACCTTTAAGGCCCCCAGGGTCCAAGCCC  496

Qy   1625 AG-ACCGGGGCTCCTAACGATGCCAAGGGGA---CTGCAAAGCCGCCCTTTTCTCAGCGGAG 1681
          ||  |||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db    497 AGAACCGGCGGCTCCTAACGATGCCAAGGGGACTGCAAAGCCGCCCTTTTCTCAGCGGAG  556

Qy   1682 AAAACCCCTTTGCCACTGTGAAAC-TCCGCCGACTGTGACGAATGATGGCTCGGCACCC 1740
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    557 AAAACCCCCTTTGCCACTGTGAAACTTCCGCCGACTGTGACGAATGATGGCTCGGCACCC  616

Qy   1741 A-TCATTCGATGAGAGGACAGCCAAGGACTCTCCCGGGCTTCCCCGGTTCTCCCTTGCGG 1799
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    617 ATTCATTCGATGAGAGGACAGCACCCAAGGACTCTCCCGGGCTTCTCCCGGTTCTCCCTTGCGG  676
```

Qy 1800 AATGATGGGGCGGCATCCTGTCTGCCACGTGGCTGAGCGGTGGGGAGAGCTTCAGTGTGGAGAGCC 1859
          |||||||||||||| |||||||||||||||| |||||| ||||||||||||||| |||||||
Db   677 AATGATGGGGCGGCGGCCTGTGTCTGCCAGGTGCGGAAGCCTTCAGTGTGGGAGAGGGC 736

Qy 1860 TAACTCTAAGTCCGCCGCTTAAGCAAATCAGGCCTGCTTAAGGCACTTCAGCGTAGTTGGGGTTGA 1919
          ||||||||| ||||||| |||||||||||||| ||| |||| |||||||||||||||||||||
Db   737 TAACTCTAAATGTCCGCCCTGCTTAAGCAAATCATGCTTGTCTGTTTCACGTAGTTGGGGTTGA 796

Qy 1920 CAAGTTTCTGCCTTTAAGATAAAATGAGTAATAGTGTAACAGCACCACCATTTAAA 1979
          |||||||||||| |||||||||||||||||| |||| ||||||||||||||||||
Db   797 CAAGTTTCTGCCTTTAAAATAAATGAGTAATAGTGTCTAATAGCCAGCCATTTAAA 856

Qy 1980 TATTTTCTTCCTATTCTGTTCAGAAACAGTAAATAGTCTCAATGTTTCAATCTTTAAAAAAAAAA 2039
          |||||||||||||||||||| ||||||||||||||||| ||||||||||||||
Db   857 TATTTTCTTCCTATTCTGTTCAAAAAACTTGGTTTCAATCTTTTAANANNNNNN 916

Qy 2040 AAAAAAAA 2047
          ||||
Db   917 NNNNNAAA 924

RESULT 14
BI335025
LOCUS   BI335025   996 bp   mRNA   linear   EST 30-JUL-2001
DEFINITION 602990031F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141190 5',
            mRNA sequence.
ACCESSION   BI335025
VERSION   BI335025.1 GI:15019682
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 996)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1347 row: n column: 07
            High quality sequence stop: 827.
FEATURES   Location/Qualifiers
     source   1..996
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="NIH_MGC_12"
                     /clone="IMAGE:5141190"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
ORIGIN

Query Match   36.3%; Score 760.4; DB 3; Length 996;
Best Local Similarity   96.5%; Pred. No. 1.3e-144;
Matches   853; Conservative   0; Mismatches 21; Indels 10; Gaps 7;

Qy   708 GGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAGAAATTCATTGCAGATGGTTGCAAGA 767
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1 GGAGACCGTTACTTCTCGTCAGAGTGAAATCCAGAAAATTCATTGCAGATGGTTGCAAGA 60

Qy   768 GGCTCTGCTTGAAGAGAAGAGGCGCCTTCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGCTCTGCTTGAAGAGAAGAGGCGCCTTCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 827

Db   61 GGCTCTGCTTGAAGAGAAGAGGCGCCTTCTTCTGCTTTCTGGTTGATAAGCACTGTGGCTTTGC 120

Qy   828 AAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTACAGAACTGAATTCCAAGCTGCCTCG 887
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   121 AAACCACATACATTATTATCACTTACAGTCTGCAGAACTACTACAGAACTGAATTCCAAGCTGCCTCG 180

Qy   888 GTGGCAGGAGACCTCGTGTTCATGCCATCAAAAGTGCCAGAGAAAATCATGGAGAAATATGATCGA 947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGGCAGGAGACCTCGTGTTCATGCCATCAAAAGTGCCAGAGAGAAAATCATCAGAATATGATCGA 240

Qy   948 AGAAATAAAGACCCCAGCCCTCTACCCCCGTGTCTCGGAACTCCTCAGGCTTCACCCATGAT 1007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AGAAATAAAGACCCCAGCCCTCTACCCCCGTGTCTCGGAACTCCTCAGGCTTCACCCATGAT 300

Qy 1008 CGAGAGAAGCAAATGTGGGTTAGGAAAGATTACGACACCCCTTTCTAAATGCTTAAATGAT 1067
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGAGAGAAGCAAATGTGGGTTAGGAAAGATTACGACACCCCTTTCTAAATGCTTAAATGAT 360

Qy 1068 GCCCCCCGCTCCTTCAGGCAGCAGCATATACCCAGTCCCTTGATCGATATGTTTAATTAACCC 1127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCCCCCGCTCCTTCAGGCAGCAGCATATACCCAGTCCCTTGATCGATATGTTTAATTAACCC 420

Qy 1128 AGCCCACGGCTG--CCCCGGAATTCACAAAAGGGTAAAATAAATTCAAACAGTTCAGAGATC 1186
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCCCACGGCTGCCCCCCGGAATTCACAAAAGGGTAAAATAAATTCAAACAGTTCAGAGATC 480

Qy 1187 CCAGTTTACAGCGATCAGTTTCCGGTTGCAACGGCTGACACATGATGAAGAAGAAGCAGAAAG 1246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCAGTTTACAGCGATCAGTTTCCGGTTGCAACGGACTGACACATGATGAAGAAGAAGCAGAAAG 540

Qy 1247 TGAAGACCATCTTCCCGCACACTGCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGG 1306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TGAAGACCATCTTCCCGCACACTGCGGGGCTCCAACAAGACCTTACTCAGCTTTGCACAGG 600

Qy 1307 GAGATGTCATCACGCGTCGCTCATCCCCGAGGAGAAGGATGGCTGGCTTATGGAGAACACG 1366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GAGATGTCATCACGCGTCGCTCATCCCCGAGGAGAAGGATGGCTGGCTTCTTATGGAGAACACG 660

Qy 1367 ACGTGTCCAAGGCGAGGGGGTTGGGTT-CCCGTCGTCGTTACACGGAAGTTGCTGGAAGAAAAT 1425
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   661 ACGTGTCCAAGGCGAGGGGGTTGGGTTTCCCGTCGTCGTTACACGGAAGTTGCTGGAAGAAAAT 720

Qy 1426 GAGA--CAGAAGCAGTGACCGTGCCCACG---CCAAGCCCCACCAGTTGAGGAAGCACATCAGC 1482
          |||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   721 GAGACCAGAAGCAGTGACCGTGCCCACGGCCCACCAGTTCGAGGAAGCACATCAGC 780

Qy 1483 ACCGTGAACTTGTCTGCAGTAGCAGTGTTGTCATCCCCCGACTACTTGGAATGC 1542
          |||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   781 ACCGTGAACTTGTCTGCAGAATAGCAGTTGTTGTCATCCCCACC---GATACTGAATGC 837

Qy 1543 TTGTCCCATGGGGGCCAGCTGCCCGACAGGA-AGCAAAATTCGGCCCG 1586
          |||||||||||||||||||||||||||||| ||||||||||||| |
Db   838 TTGT-CATGGGGGGCCAGCTGCCCGACAGGA-AGCAAAATTCGGCCCCG 879

RESULT 15
BG761459
LOCUS   BG761459   933 bp   mRNA   linear   EST 15-MAY-2001
DEFINITION 602718666F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858696 5',
            mRNA sequence.
ACCESSION   BG761459
VERSION   BG761459.1 GI:14072112
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)

```
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1713  row: c  column: 17
             High quality sequence stop: 851.
FEATURES             Location/Qualifiers
     source          1..933
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4858696"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_49"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

             Query Match
             Best Local Similarity   36.1%;
             Matches  885;  Conservative    0;  Mismatches   28;  Indels   11;  Gaps   10;
```

```
Db   542  ATGGGGGCAGCTGCCGACAGGAGGAGGAGGAGCAGAGATTCGGCGGCAAAGCGGGGACGGGGGGCTCCTTAACGATGCCAAAAGCCCGCC  601
Qy  1609  CCAGCGTCCAAG-CCCGAGCACGGGGGGCTCCTTAACGATGCCAAAAGCCCGCC  1667
Db   602  CCAGCGGTCCAAGTCCCGAGACCCGGGCGCTCCTTAACGATGCCAAACGGGGACGGGGGACTGCAAAGCCGCC  661

Qy  1668  TTTTCTCAGCGGGAGGAGAAAACCCCTTTGCCACTGTGAAACTCCGCCCGACTGTGACGAATGA  1727
Db   662  TTTTCTCAGCGGGAGGAGAAAACCCCTTTGCCACTGTGAAACTCCGTCGGACTGTGACGAATGA  721

Qy  1728  TCGGCTCGGGCACCCATCATTCGATCGAGGAGGACAGCCAAGGACTCTCCCGGGCCTCTCCGGT  1787
Db   722  TCGGCTCAGCACCCATCATTCGATGAGGAGGACAGCC-AGGACTCTCCCGGGGGCTTCTCCGGTT  780

Qy  1788  TCTCCCTTCCGGAATGATGGGGCCGCATCCTGTC-TGCCACGTGCTGACGGTCGGGGAAGCTT  1846
Db   781  CTCCCTTGCGGAATGATGGGGCGCGGTCCTTCTTCCACGTTGCTGACGGTCGGGGAAGCCT  840

Qy  1847  CAGTGGA--GAGGCCTAACTCTAAATGTC-GCCTCGCTTAAGCAAAT-CATGCTTCTCTGTT  1902
Db   841  CAGTGGGAAGAGGGGCCTAACTCTTAAGTCGGCCTCGCTTAAGCAAATCCAATCCATGCTTCTCTGTT  900

Qy  1903  TCACGTAGTTGGGTTGACAAGTTT  1926
Db   901  TCCCGTAGTTGGGTTGCCAGTTCT  924

Qy  1013  GAAGCAATGTGGTTAGGAAAAGATTACGACACACCCTTTGTAAATGCTCACCAAAGATGCCC  1072
Db     2  GAAGCAATGTGGTTAGGAAAAGATTACGACACACCCTTTCTAAATGCTCACCAAAGATGCCC  61

Qy  1073  CCGCTCCTTCAGCGACGAGCAATATACCAGTCCCTTCCTTGATCGATATGTTTAATAACCCAGCCA  1132
Db    62  CCGCTCCTTCAGCGACGAGCAATATACCAGTCCCTTGATCGATATGTTTAATAACCCAGCCA  121

Qy  1133  CGGCTCCCCCGAAATTCACAAAGGGTAAATATAATTCACACAGGTTACTTCCGGAAGATCCCAGTT  1192
Db   122  CGGCTGCCCCGAAATTCACAAAGGGTAAATAATTCACACAGGTTACTTCCGAAGATCCCAGTT  181

Qy  1193  TACAGGCCATCAGTTTCCGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA  1252
Db   182  TACAGGGATCAGTTTCGGTTGCAACGGGACTGAACATGATGAAGAAGCAGAAAGTGAAGA  241

Qy  1253  CCATCATCTTCCCGC-ACACTGCCGGGCCTCCAA-CAAGACCTTACTTACTCAGCTTTGCACAGGGAGA  1310
Db   242  CCATCTTCCCGGACAACTGCGGGCCTCCAAGCAACACCTTACTCAGCTTTGCACAGGGAGA  301

Qy  1311  TGTCATCACGCTGCTCATCCCCGAGCAGAGACCGGATGGCTGGCTCTTACTGGAGAGAACACGACGT  1370
Db   302  TGTCATCACGCTGCTCATCCCCGAGCAGAGACCGGATGGCTGGCTCTTACTGGAGAGAACACGACGT  361

Qy  1371  GTCCAAGGCGAGGGGGTTGGGTTCCCGTCGTCGTCGTACCGAAGTTGCTGGAAGAAAATGAGAC  1430
Db   362  GTCCAAGGCGAGGGGGTTGGGTTCCCGTCGTCGTCGTACGGAAGTTGCTGGAAGAAAATGAGAC  421

Qy  1431  AGAAGCAGTGACC-GTGCCACGCCAAGCCCCA-CACCAGTGAGAGAGCATCAGCACCGTG  1488
Db   422  AGAAGCAGTGACCTGTGCCCACGCCAAGCCCCAGCACCAGTGAGAGAGCATCAGCACCGTG  481

Qy  1489  AACTTGTCTCAGAAATAGCACAGTGTTGTCATCCCCCCACCCGAACTGGAATGCTTGTCC  1548
Db   482  AACTTGTCTGAGAATAGCAGTGTTGTCATCCCCCCACCCGACTACTTGGAATGCTTATCC  541

Qy  1549  ATGGGGGCAGCTGCCGACAGGAGGAGGAGGAGCAGAGATTCGGCGGCAAAGCGGGGACGGGGGGCTCCTTAACGATGCCAAAAGCCC  1608
Db             ||||||||||||||||||||||||||||||||||||||||||||||
```

```
            Score 757.2;  DB 2;  Length 933;
            Pred. No. 6e-144;
```